

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:41:32 ; Search time 8423 Seconds
(without alignments)

10993.468 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacgaggagagaccata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	6 AR533623	AF533623 Sequence
2	916.8	56.3	1705	15 AF452450	AF452450 Glycine m
3	904	55.5	1795	6 AR533622	AF533622 Sequence
4	718.4	44.1	1648	15 AY353089	AY353089 Populus x
5	706.8	43.4	1753	15 CRU63784	U63784 Catharanthu
6	664.8	40.8	1835	15 BT013254	BT013254 Lycopersi
7	663.8	40.7	1386	15 AY568717	AY568717 Lycopersi
8	663.6	40.7	1608	15 SCJ001208	AJ001208 Brassica
9	659.4	40.5	103129	15 AP004494	AP004494 Lotus cor
10	657.2	40.3	1377	15 AY103313	AY103313 Arabidops
11	657.2	40.3	1555	15 AY093319	AY093319 Arabidops
12	657.2	40.3	1654	15 AY062665	AY062665 Arabidops
13	657.2	40.3	1702	15 AY087591	AY087591 Arabidops
14	657.2	40.3	1709	15 AF428445	AF428445 Arabidops
15	657.2	40.3	1712	15 AY054175	AY054175 Arabidops
16	656.6	40.3	1710	15 AF424582	AF424582 Arabidops
17	655	40.2	1714	15 AY088011	AY088011 Arabidops
18	653.2	40.1	1398	15 BT002612	BT002612 Arabidops

19	652.4	40.0	1763	15 ATU53865	US3865 Arabidopsis
20	651.6	40.0	1398	6 CQ805882	CQ805882 Sequence
21	651.6	40.0	1398	6 AX412862	AX412862 Sequence
22	651.6	40.0	1398	6 AX505803	AX505803 Sequence
23	651.6	40.0	1398	6 AX651565	AX651565 Sequence
24	650.2	39.9	1748	15 ATU53864	US3864 Arabidopsis
25	620.8	38.1	1482	15 ATU56922	US6922 Arabidopsis
26	615.8	37.8	1634	15 ATU43412	U43412 Arabidopsis
27	602.8	37.0	73519	14 AP007975	AP007975 Lotus cor
28	594.2	36.5	1700	15 SCJ001207	AJ001207 Brassica
29	588.2	36.1	1687	15 AY088665	AY088665 Arabidops
30	586.6	36.0	1712	15 ATU96045	U96045 Arabidopsis
31	578.4	35.5	1783	15 AF360192	AF360192 Arabidops
32	577.4	35.4	1365	6 CQ805458	CQ805458 Sequence
33	577.4	35.4	1365	6 AX412892	AX412892 Sequence
34	577.4	35.4	1365	6 AX412893	AX412893 Sequence
35	577.4	35.4	1365	6 AX507610	AX507610 Sequence
36	576.6	35.4	1365	15 AY040005	AY040005 Arabidops
37	576.4	35.4	1365	15 AF023167	AF023167 Arabidops
38	570.6	35.0	1738	15 ATU53866	US3866 Arabidopsis
39	568	34.9	1945	15 AF016282	AF016282 Arabidops
40	568	34.9	47319	6 AX059533	AX059533 Sequence
41	568	34.9	103632	15 AF074021	AF074021 Arabidops
42	568	34.9	200001	15 ATCHRIV13	AL161501 Arabidops
43	562.4	34.5	1450	15 ATU56921	US6921 Arabidopsis
44	561.6	34.5	2945	15 AF016284	AF016284 Arabidops
45	561.6	34.5	103270	15 ATF1N20	AL022140 Arabidops

ALIGNMENTS

RESULT 1	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	AR533623					
DEFINITION	AR533623					
ACCESSION	AR533623					
VERSION	AR533623.1	GI:53923592				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1629)					
AUTHORS	Falco,S.C.					
TITLE	Genes encoding plant adenosine 5'-phosphosulfate reductase					
JOURNAL	Patent: US 6730827-A 7 04-MAY-2004;					
FEATURES	E. I. du Pont de Nemours and Company; Wilmington, DE					
source	Location/Qualifiers					
	1..1629					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN

Query Match	100.0%;	Score 1629;	DB 6;	Length 1629;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1629;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCATCTTCAATTC	60	
Db	1	GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCATCTTCAATTC	60	
Qy	61	CGACCAACTTCCACCTTCCCATCGGAACCCAACTTCCGCAATTTGGGTCAATAG	120	
Db	61	CGACCAACTTCCACCTTCCCATCGGAACCCAACTTCCGCAATTTGGGTCAATAG	120	
Qy	121	GATTCGGAGAGGCCCATTCGGAGCGCGGTAAATTTCAATTTATCTCAAGACGAGCTT	180	
Db	121	GATTCGGAGAGGCCCATTCGGAGCGCGGTAAATTTCAATTTATCTCAAGACGAGCTT	180	
Qy	181	GGTAAAGCCGTTAAACGCGAACCTCCACGAGGATTCCTCTCCGAGCAAC	240	
Db	181	GGTAAAGCCGTTAAACGCGAACCTCCACGAGGATTCCTCTCTCGAGCAAC	240	


```
Qy 41 GCTTTCACCTTCTTCAATTTCCGACCAACTTCCACCTTCCATCATCGGAACCCAAACTT 100
Db 55 GCTGCTGACCTTCTAGCTCTTCTCTCTCGATCTTCTCTCGATCTTCTCTCGACGCAAAAGCT 114
Qy 101 CCGCAAAATGGGTCAATTTAGGATTTCCGAGAGGCCATTTGGAGCGCGTAAATTTCAAT 160
Db 115 CCTCAATTCGGTTTCTGTTTCCGAGAGGTCGCTTGTCTGTTGTTGTTGTTGTTGTTCAAT 174
Qy 161 TTATCTCAAGACGGAGCTTGGTAAGCCGTTTAAACCGCCGAACCTCCACGCAAGGATTC 220
Db 175 GTAACCTCAAGACGCTCTCTGTTGAGGCCACTCAACGCGGAACCGCAACGAACTCT 234
Qy 221 ATTGTTCTCTGCGAGCAACCAACATCTGTTGCTTCTGCTTCTGAGAGCAAGAGGAAGAT 280
Db 235 ATTGTTCTCTGCGAG---CAACTATCTGTTGCTCTGAGGTTGAGAGGAAGAGGAAGAT 291
Qy 281 TTTGAACAGATAGCGAGTATCTGACAAATGCTTTTCACTCTTGAATCATGATAGAGGCC 340
Db 292 TTTGAGCAATAGCGAAGACCTTGAANAATGCATCTCTCTTGAANAATATGATAGAGGCC 351
Qy 341 CTCGACAAATTCGGCAACGACATAGCTATTGSCCTTCAGTGTGCTGAAGATGTTGCTTTG 400
Db 352 CTCGAGAAATTTGGGAACGACATCGCTATTGCAATTAGTGTGCTGAAGATGTTGCTTTG 411
Qy 401 ATTGAGTATGCGAAATTTGAACGGTTCGACCCCTTTTAGGGTTTTCAGTTTGGACATCTGGAGA 460
Db 412 ATTGAGTATGACATTTTGAACGGTTCGACCCCTTACAGAGTGTGTTAGTCTTGAACATCTGGAGA 471
Qy 461 CTGAACCCAGAACTTATCACTTTTGTGATGCGGTTGAGAGCAATTTGGAATTCGCATT 520
Db 472 CTGAATCTGAHAACCTTCAAAATTTTGTGCGGTTGAGAGCAATTTGGAATTTCACTT 531
Qy 521 GAGTACATGTTTCCCTGATGCTGTTGAGGTTTCAAGCATTTGTTGAGGATTAAGGGGTTATTC 580
Db 532 GAGTACATGTTTCCCTGATGCGTTGAGGTTTCAAGCATTTAGTAGAACTTAAGGGCTTTC 591
Qy 581 TCTTTTACAGAGATGGGACCAAGAGTGTGTCAGGTTGAGAAAGTGTGAGGCTTTTAAG 640
Db 592 TCATTTTACAGAGATGGGACATCAAGAGTGTGTCGAGTAAGGAGTGTGAGGCCCTTGAG 651
Qy 641 AGGSCCTTAAAGGTCCTCAGAGCATGATAACTGTTGAGAGGAAGACCATGTCACCTGTT 700
Db 652 AGAGCCCTTAAAGGTCCTCAAGAGCATGATCACTGACAGAGAAAGATCACTCTCTGTT 711
Qy 701 ACTAGTCTGAATATACCGGTTGTTTCAAGTGTGATCCGCTTTTGAAGGAATGGATGGGA 760
Db 712 ACTAGTCTGAATATCCCTGTTGTCAGGTTGATCTGTTTGAAGGACTGGATGGGA 771
Qy 761 ATTGGAAGCTTGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGAAC 820
Db 772 ATTGGCAGCCTGTGTAAGTGAACCCCGTTGCAAAATGTTAATGTTCTAGACATATGGAGC 831
Qy 821 TTCTTAGGACCATGATGTGCTGCAATTTCTTGCATGCAAAAGATATGTTTCCATT 880
Db 832 TTCTTAGGACCATGATGTGCTGCAATTTCTTGCATGCAAAAGATATGTTTCCATT 891
Qy 881 GGGTGTGAGCCTGCACTAGGCTGTTTAACTTGGGCAACATGAAGGAAGGAGGTGG 940
Db 892 GGGTGTGAGCCTGCACTAGGCTGTTTAACTTGGGCAACATGAAGGAAGGAGGTGG 951
Qy 941 TGGTGGGAGGATGCCAAGCTAAGGAATGTTGTTTCCAAAGGAATATGAAGGAGCAG 1000
Db 952 TGGTGGGAGGATGCCAAGCTAAGGAATGTTGTTTCCAAAGGAATATGAAGGAGCAG 1011
Qy 1001 AAAGAGGAGATGTTAATGAATAGGCTATCCCAATCCCAATGGAATGTTGATGCTACC 1060
Db 1012 GATGCTGCCAGCTTAAATGAATAGGCTATCCCAATGGAATGGA-----TCTGCC 1062
Qy 1061 ACTGTGCTGACATTTTCAACAGCCGAATGTTAGTTAACTTTGAGCAGGACTGGAATTTGAG 1120
Db 1063 ACTGTGCTGACATTTTCAACAGCCGAATGTTAGTTAACTTTGAGCAGGCTGGAATTTGAG 1122
Qy 1121 AATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGCTTTATGCAACCATGG 1180
```

```
Db 1123 AATTTGGCAAAATTTGGAGAACCGAAAGAACCCCTGGCTTGTGTTGCTCTATGCAACATGG 1182
Qy 1181 TGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGACAGTATTAGCAGG 1240
Db 1183 TGCCGCTTCTGTCAAGGCTATGGAGGAATCATATGTTGATTGTTGGCAGAAAAGTTAGCTGG 1242
Qy 1241 TCACACGGGATGAAGGTTGGAAAATTTTAGAGCAGATGGAGAACAGAAAGAAATTTGCAAG 1300
Db 1243 T---CAGGAGTGAAGTTGCAAAATTCAGAGCGATGGAGACAGAGGAATATGCAAG 1299
Qy 1301 AGTGAATCTGAATTTGGGAAGCTTCCCTACGATATATTTTCCAAAGCATTCGCTCGG 1360
Db 1300 ACTGAATCTGAGTTGGGAAGCTTCCCAACAATACTACTCTTCTCCCAACAACCTTCT--- 1356
Qy 1361 CCAACATAAAGTATCTCTCAGAAAAGAGAGATGTTGATTTCTTGTGATGGCATTTGTAAT 1420
Db 1357 CAACCAATTAAGTACCTTCGAAAAGAGAGATGTTGATTTCAATTGACGGCATTCGTGAAT 1416
Qy 1421 GCCTTAAGATGAGGAT 1436
Db 1417 GCTTACGGTGATGGT 1432

RESULT 3
AR533622 LOCUS AR533622 1795 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6730827.
ACCESSION AR533622
VERSION AR533622.1 GI:53923591
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1795)
AUTHORS Falco, S.C.
TITLE Genes encoding plant adenosine 5'-phosphosulfate reductase
JOURNAL Patent: US 6730827-A 5 04-MAY-2004;
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
source 1..1795
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 55.5%; Score 904; DB 6; Length 1795;
Best Local Similarity 81.3%; Pred. No. 2e-227;
Matches 1103; Conservative 0; Mismatches 235; Indels 18; Gaps 4;
Qy 81 CATCATCGGAACCCAACTTCCGCAAAATTCGGTCAATTAGGATTTCCGAGAGGCCCATTTG 140
Db 135 CTTTCATCGGACGCTAAAGCTCCGCAAAATTCGTTCCCTTCGGTTTCCGAGAGGCCCTCAAG 194
Qy 141 GAGGCGCGTTAATTTCAATTTATCTCAAAAGAGGAGCTTTGGTAAAGCCCGTTTAACGCCG 200
Db 195 TTTTCGCTGCTGTTGTTTAAATTTAACTCAAAAGAGCGCTCTCGGTGAGGCCCATCAATGCCG 254
Qy 201 AACCTCCAGCAGAGGATTCATTTGTTCCCTCTCCGAGCAACCAACCATGCTTGTCTCTGTT 260
Db 255 AACCGCAACGGAATGATTTGTTGTTTCTCTCTTGAG---CAACTATCGTTGCTCTGAGG 311
Qy 261 CTGAGACGAAAGAGGAAGATTTTGAACAGATAGCAGTATCTCGACAAATGCTTTCACCTC 320
Db 312 TTGAGAGGAGGAAGAAGATTTTGAAGCAATTAGGGAAGACCTTGAANAATTCATCTCTC 371
Qy 321 TTGAATATCATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTCTCAGTG 380
Db 372 TTGAGATTTAGGATAAGGCCCTCGAGAAAATTTGGGAACGACATCGCTATTGCTCTTAGTG 431
Qy 381 GTCTGAGATGTTGCTTTGATTTGATGATCGAAATTTGACGGGTGCAACCTTTAGGCTTT 440
Db 432 GTCTGAGATGTTGCTTTGATTTGATGATGATGATTTGACGGGTGCAACCTTTACAGAGTGT 491
```

```
Qy 441 TCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTGCGGTTGAGA 500
Db 492 TTAGTCTTGACACTGGGAGACTGAACCCAGAACTTATCAAAATTTTGTGACGCTGTTGAGA 551
Qy 501 AGCAATTATGGAATTCGCAATGAGTACATGTTCCCTGATGCTGTGTAGGTTTTCAGGCAATGG 560
Db 552 AGCAATTATGGAATTCATTTAGTACATGTTCCCTGATGCGTTGAGGTTTCAGGCAATAG 611
Qy 561 TGAGGAGTAAGGGGTTATCTCTTTCTACGAGGATGGCCACCAAGAGTGTGTGACGGTGA 620
Db 612 TAAGAACTAAGGGGCTCTTCTCATTTTACGAGGATGGCCATCAAGAGTGTGTGAGAGTAA 671
Qy 621 GAAAGGTGAGGCCCTTTAAGGAGGCCCTTAAGGCTCTCAGAGCATGATACTGGTCAGA 680
Db 672 GAAAGGTGAGGCCCTTCAGGAGAGCCCTTAAGGCTCTCAAGGATGATTAAGGATGAGTGA 731
Qy 681 GGAAGACCACTGACCTGCTGCTAGGCTCTGAAATACCGGTTGTTTTCAGGTTGATCCGCTT 740
Db 732 GAAAGACCACTGCTGCTGCTAGGCTCTGAAATACCGGTTGTTTTCAGGTTGATCCGCTT 791
Qy 741 TTGAGGGAATGGATGGTGGAAATGGAAGCTTTGGTGAAGTGGAAACCCCTGTTGCAAAATGTA 800
Db 792 TTGAGGGAATGGATGGTGGAAATGGAAGCTTTGGTGAAGTGGAAACCCCTGTTGCAAAATGTA 851
Qy 801 AGGCCATGACATATGGAATCTCTTAGGACCATGAATGCTGCTGTAATTCCTTGGCATG 860
Db 852 ATGCTCTAGACATATGGAATCTCTTAGGACCATGAATGCTGCTGTAATTCCTTGGCATG 911
Qy 861 CAAAGGATATGTTTCCATTTGGTGTGAGCCCTGACCTAGGCTGTTTTCCTTGGGCAAC 920
Db 912 CCAAGGATATGTTTCCATTTGGTGTGAGCCCTGACCTAGGCTGTTTTCCTTGGGCAAC 971
Qy 921 ATGAAAGGGAAGGAGGTGGTGGGAGGATGCCAAAGCTTAAGAAATGTTGGTCTTTCACA 980
Db 972 ATGAAAGGGAAGGAGGTGGTGGGAGGATGCCAAAGCTTAAGAAATGTTGGTCTTTCACA 1031
Qy 981 AAGGAATATGAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCC 1040
Db 1032 AAGGTAATTTGAAACAGGAAGATGCTGCCAGCTTAATGGAATGGGACCTCCCAAGGAA 1091
Qy 1041 ATGCAATATGGTGTATGCTACCTGCTGCTGACATTTTCAACAGCCCGCAATGTAGTTAACT 1100
Db 1092 A-----TGGCTCTGCCACTGTGCTGACATTTTTCATCTCCAGAAATGTGTGAGCT 1142
Qy 1101 TGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGAACCATGGCTTG 1160
Db 1143 TGAGCAGGCTCCGGATTCAGAAATTTGGCAAAATTTAGAGAACCGAAAGAACCACTGGCTTG 1202
Qy 1161 TTGTGCTTTATGACCATGGTGGCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACT 1220
Db 1203 TTGTGCTCTATGACCATGGTGGCCCTACTGCCAGGCTATGGAGGATCGTATGTTGATC 1262
Qy 1221 TAGCAGACAAATTTAGCAGGGTCAACAGGAGATGAGGTTTGGAAATTTAGAGCAGATGAG 1280
Db 1263 TGCCAGAGAAATTTAGCAGGTT---CAGAGTGAAGTTTGCAAAATTCAGAGCCGATGAG 1319
Qy 1281 AACGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAGGCTTCCCTACGATATATTTT 1340
Db 1320 AGCAGAGGAATATGCAAGAGTGAATGCAATTTGGGAGGCTTCCCTACGATATATTTT 1379
Qy 1341 TCCCAAGCAATTCGCTCGGCCCAACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATT 1400
Db 1380 TCCCAAGCAATTCGCTCGGCCCAACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATT 1436
Qy 1401 CTTTCATGCGCATTTCTTAATGCTTTAAGATGAGGAT 1436
Db 1437 CATTGACGCGCATTCGTGAATGCCCTTACGTTAGTGT 1472
```

RESULT 4
AY353089 1648 bp mRNA linear PLN 12-JAN-2005
LOCUS AY353089
DEFINITION Populus x canescens adenosine 5' phosphosulfate reductase mRNA,

```
complete cds.
AY353089
AY353089.1 GI:34099826
.
Populus alba x Populus tremula
Populus alba x Populus tremula
Bukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 1648)
Kopriva,S., Hartmann,T., Massaro,G., Hoenicke,P. and Rennenberg,H.
Regulation of sulfate assimilation by nitrogen and sulfur nutrition
in poplar trees
Trees (Berl. West) 18, 320-326 (2004)
2 (bases 1 to 1648)
Hartmann,T., Hoenicke,P., Wirtz,M., Hell,R., Rennenberg,H. and
Kopriva,S.
Regulation of sulphate assimilation by glutathione in poplars
(Populus tremula x P. alba) of wild type and overexpressing
gamma-glutamylcysteine synthetase in the cytosol
J. Exp. Bot. 55 (398), 837-845 (2004)
3 (bases 1 to 1648)
Kopriva,S.
Direct Submission
Submitted (28-JUL-2003) Institute of Tree Physiology,
Georges-Koehler-Allee 053, Freiburg 79085, Germany
Location/Qualifiers
FEATURES
source
1..1648
/organism="Populus alba x Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:80863"
/clone="INRA 717 1B4"
74..1471
/codon_start=1
/product="adenosine 5' phosphosulfate reductase"
/db_xref="GI:34099827"
/translacion="MALVSSSSSISASGFSRSSQELKAPQFGSKLLDRPTTFVNV
SQRCVAKPLNAPKRNQSVVPLAATAPEIAEKVEVEDEYKLAKELVNGSPLEIM
DKALEFGDDIAIAFGSAGDAIEYAKLTGPERFVSLDTGRNLNPTTFHFDQVEKH
YGRLEYMFPDAVEQALVRNKLGSFYEDGHQCCRVKRVPLRLRALKGLRAWITGQ
RKQSPCTSEIPIVQVDPVPEGLDGGAGSLIKWNPANVEGQDVWPKFLRTMDLPVNS
LHSGYISIGCEPCTRPVLQOHEREGRWMDATAKECGLHKNLKGDDAOLNGNG
NGAHANGAATVADIFNSNLVNLSPGIEINLLKLNKRPWLVLVAPWCQFCQGM
ASYVELADKLKLAGSVKVGKFRADGDKGFSKQELQLGFFPILFPKHSSRPKYPSE
KRDVDSLMTFVNLR"
ORIGIN
Query Match 44.1%; Score 718.4; DB 15; Length 1648;
Best Local Similarity 72.3%; Pred. No. 2.1e-178; Indels 21; Gaps 5;
Matches 1009; Conservative 0; Mismatches 366;
Qy 43 TTTCACTCTTCAATTTCCGCAACCAATTCACCTTCCCATCATCGGAACCAAACTTCC 102
Db 91 TTCTTCTTCTCGATTTCTGCTTCTGGGTTTTCAGGTTCTAGTCAAGAACTTAAAGCACC 150
Qy 103 GCAAAATGGGTCAATTAGGATTTCCGAGAGGCCCATTCGAGAGCCGCTTAATTTCAATTT 162
Db 151 CCAGTTTGGTCTTTTAACTGTTTAGATCGGCCGA-----CAACATTTGTTAATGTTTC 204
Qy 163 ATCTCAAGAGCGAGCTTGGTAAAGCCGTTAAACGCCGAACCTCCACCAAGGATTCAT 222
Db 205 TTCAAAAGAGCGCTGTGCTGTAAAGCCCTTTAAATGCTGAGCCTTAAAGAAATGTTCTGT 264
Qy 223 TGTTCCTCTCGCAGCAACCAACCATCGTTGCTTCTGCTTCTGAGACGAAA-----GAGA 276
Db 265 TGTTCCTCTCGTCAACTATCGCTGCTCCAGAGATTCAGAGAAAGTAGAGGTGAGGA 324
Qy 277 AGATTTTGAACAGATAGCCAGTGTCTCGACAAATGCTTCACTCTTGAATCATGATAG 336
Db 325 AGATTTTGAAGATTAGCTAAGGAGCTTGTAAATGGTTCTCTCTTGGAAATTTATGATAA 384
```



```
QY 337 AGCCTCGACAAATTCGGCAACGACATAGCTATTATGCTTTCAGTGGTGTCTGCTGAAGATGTGC 396
Db 385 GGCATTTGAGAAATTTGGGATGACATTTGCTATTGCTTTTTCAGTGGTGTCTGCTGAAGATGTGC 444
QY 397 TTTGATTGAGTATGCGAAATTTGCGGTGCGACCTTTAGGGTTTTCAGTTTGGACATGG 456
Db 445 TTTGATTGAGTATGCGAAATTTGCGGTGCGACCTTTAGGGTTTTCAGTTTGGACATGG 504
QY 457 GAGACTGAACCCGAAACTTTATCAACTTTTGTGATGCGGTGAGAAAGCAATTATGGAATTCG 516
Db 505 GCGTTGAACCCGAGACATACCAATTTCTTTGACCAAGTTGAGAAGCACTATGCAATCG 564
QY 517 CATGAGTACATGTTCCCTGATGCTGTGTAGGTTTCAGGCAATTTGAGAGGATGAAGGGTT 576
Db 565 AATTGAGTACATGTTTCCAGATGCTGTGTAGGTTTCAAGCAATTTAGTAAGGAACAAGGGTTT 624
QY 577 ATTCTCTTTTACGAGATGGGCAACCAAGAGTGTGAGGTTGAGAAAGTGTAGGCTTTT 636
Db 625 ATTCTCTTTTACGAGATGGGCAATCAGGATGCTGCTGCTGAGGAGGTTGAGACCTTT 684
QY 637 AAGGAGGCGCTTTAAGGGTCTCAGAGCATGATAAATGCTGAGAGGAAGACCAAGTCAAC 696
Db 685 GAGCGGCTCTGAAGGGCTTCGGGCTGATCACTGGCCAAAGGAAGATCAATCTCC 744
QY 697 TGGTACTAGGTCTGAATAACCGGTGTTTCAGGTTTTCAGGCTTTTTCAGGGAATGAGTGG 756
Db 745 AGGAGCAGGTCTGAATAATCCAGTTGTTTCAGGTTGAGACCGCTTTTTCAGGGAATGAGTGG 804
QY 757 TGGAAATTTGGAAGCTTGTGAGTGAAGTGAACCTGTTGCAAAATGTAAGGCGCATGATG 816
Db 805 TGGGCTTGAAGCTTATCAAGTGAATCCAAATGGAATGTTGAGGGAACAAGATGATG 864
QY 817 GAACCTCTTTAGGACCATGAATGTGCTGTGAATTCCTGATGCAAAAGCAATGTTTC 876
Db 865 GAAGTTCTTTGCAACTATGATGTGCTGTGAATTCATGCAATCAAGGGAATGATCTC 924
QY 877 CATGGGTGAGCCTGCACTAGGCTGTTTACCTGGGCAACATGAAGGAAGGAG 936
Db 925 TATTGGTGGAGCTTTGCAAGGCCAGTCTCTCTGGCCAAACATGAGAGAGAAGGACG 984
QY 937 GTGGTGTGGAGGATGCCAAGCTAAGGATGAGTGTGCTTTCACAAAGGAATGTAAGCA 996
Db 985 GTGGTGTGGAGGATGCCAAGCTAAGGATGAGTGTGCTTTCACAAAGGAATGTAAGCA 1044
QY 997 GCAGAAAGAGGAGATGTTAATGGAATGGGCTATCCCAATCCATGCAAAATGGTGTATGC 1056
Db 1045 GGGTATGAGCCCAACTTAATGGCAATGGAACGGGGCTGCCCATGCAATGG---GGC 1101
QY 1057 TACACTGTGCTGACATTTTCAACAGCCGCAATGTAGTTAACTTGAGCAGGACTGGAAT 1116
Db 1102 TGCCACTGTGCTGATATTTTCAACTCCGAGAAATTTGGTCAATTTGAGCAGGCCCGGAT 1161
QY 1117 TGAGAAATTTGGCAAAATTTGAGGACCGAAGGACCAATGCTTGTGTGCTTTATGCACC 1176
Db 1162 CGAGAACTTGTAAATTTAGAGAACCGAAGGACCAATGCTTGTGTGCTTCTACGCCCC 1221
QY 1177 ATGGTGGCCCTACTGCGAGGCTATGAGGAATCTTATGTTGACTTAGCAGACAAAGTTAGC 1236
Db 1222 ATGGTGCCAAATTTGCAAGGATGAGGACATCTTATGTTGAATTTGGCCATGAGTTAGC 1281
QY 1237 AGGGTCAACAGGATGAAGTTTGAAATTTAGAGCAGATGAGGAGAACAGAAAGAAATTCG 1296
Db 1282 AGGG---AGTGGAGTAAGGTGGGAAATTTGAGGCGAGATGGCGATCAGAAGGGATTC 1338
QY 1297 AAAGAGTGAATGGAAGTTCCTTACGATATTTATTTTCCCAAGCAATTCGTC 1356
Db 1339 AAACAGGAATTAAGTGAAGTTCCTTACGATATTTATTTTCCCAAGCAATTCGTC 1398
QY 1357 TCGGCCAAACAAGATGATCCCTCAGAAAAAGAGATGTTGATTCCTTGATGCAATTTGT 1416
Db 1399 TCGCC---CAATCAAGTACCCCTCTGAGAAAGGAGGCTGATTCATTTGATGACTTTGT 1455
QY 1417 AATGCTTAAAGTGA 1432
```

```
Db 1456 GAATGCTCTCCGGTAA 1471
|||||
RESULT 5
CRU63784
LOCUS
DEFINITION
Catharanthus roseus PAPS-reductase-like protein (par2neu) mRNA,
complete cds; nuclear gene for chloroplast product.
ACCESSION
U63784
VERSION
U63784.2
GI:12831473
Catharanthus roseus (Madagascar periwinkle)
Catharanthus roseus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vincaceae; Catharanthus.
1 (bases 1 to 1753)
Prior,A., Uhrig,J.F., Heins,L., Wiesmann,A., Lillig,C.H.,
Stoltze,C., Soll,J. and Schwenn,J.D.
Structural and kinetic properties of adenylyl sulfate reductase
from Catharanthus roseus cell cultures
Biochim. Biophys. Acta 1430 (1), 25-38 (1999)
10082930
2 (bases 1 to 1753)
Uhrig,J.F., Prior,A. and Schwenn,J.D.
Direct Submission
Submitted (15-JUL-1996) Biochemistry of Plants,
Ruhr-Universitaet-Bochum, Universitaetstr. 150, Bochum 44780,
Germany
3 (bases 1 to 1753)
Prior,A. and Schwenn,J.D.
Direct Submission
Submitted (03-FEB-2000) Biochemistry of Plants,
Ruhr-Universitaet-Bochum, Bochum 44780, Germany
Sequence update by submitter
On Feb 15, 2001 this sequence version replaced gi:1488042.
Location/Qualifiers
1..1753
/organism="Catharanthus roseus"
/mol_type="mRNA"
/db_xref="taxon:4058"
1..1753
/genes="par2neu"
/evidence=experimental
106..1500
/genes="par2neu"
/function="reduction of adenylyl sulfate (APS)"
/notes="C-terminus similar to pdi and beta-subunit of
4-hydroxylase, PDB Accession Number: 1MEK"
/codon_start=1
/evidence=experimental
/product="PAPS-reductase-like protein"
/protein_id="AAB05871.2"
/db_xref="GI:12831474"
translation="MALFTSTAIHGLSSSPQTKAAAOFGSPOLDPHTTSPS
VNVSRPLKAPINAEPKNEISIVPSAATTVAPESEKVDVEDYKLADELQNASPLE
IMDKSLAKGNDIAIAFSGAEDVALLEYAHLTGPRPRVSLDTRLNPEYKFPDIVE
KYQTHIEYMFDAVEQALVRSGILFYSFYBDHGQCCRRVKRVRPURLALKURAWIT
GQRKQSPGTRSEIPIVQVDVPFEGDGGVSLKWNPNVANVEGKDIWNFLRAMDVPV
NTLHSGQYVSIIGCEPTRPVLQGHREGRWMMEDAKACEGLHKGNIKEETLNNGN
GAVNGSDTIADIPTDNTVTSLSRGIENLLKBERREALVVLVAPWCRFCQAMEG
SVLEAEKLAGSGVKFKADGDQKAFQAQELQLNSFTLILFFPKHSSKPIKYPSEK
RVDLSLMAFVNALR"
transit_peptide 106..282
variation
1027..1030
/genes="par2neu"
/genes="par2neu"
/notes="results in cysteine to tryptophan substitution"
/citation={1}
/replaces="tgt"
1072..1075
variation
```

```
/gene="par2neu"  
/note="results in glutamine to aspartic acid substitution"  
/citation=[1]  
/replace="gac"  
1084..1086  
/gene="par2neu"  
/note="results in glutamic acid to glycine substitution"  
/citation=[1]  
/replace="gga"  
1093..1095  
/gene="par2neu"  
/note="results in asparagine to isoleucine substitution"  
/citation=[1]  
/replace="ata"  
1096..1098  
/gene="par2neu"  
/note="results in asparagine to isoleucine substitution"  
/citation=[1]  
/replace="ata"  
1099..1101  
/gene="par2neu"  
/note="results in asparagine to tryptophan substitution"  
/citation=[1]  
/replace="tgg"  
1102..1104  
/gene="par2neu"  
/note="frameshift results in glycine to tryptophan substitution"  
/citation=[1]  
/replace="tgg"  
1105..1107  
/gene="par2neu"  
/note="results in asparagine to aspartic acid substitution"  
/citation=[1]  
/replace="gat"
```

ORIGIN

```
Query Match 43.4%; Score 706.8; DB 15; Length 1753;  
Best Local Similarity 71.9%; Pred. No. 2.5e-175;  
Matches 1017; Conservative 0; Mismatches 367; Indels 30; Gaps 6;
```

```
QY 31 AATGGCCCTCGCTTCACTTCTCAATTC-----CGCACCAACTTCCACTTCCC 81  
DB 105 AATGGCTTGGCTTCACTTCTCAACTGCAATTCATGGCTCTCTGCTTCTTCAATTGA 164  
QY 82 ATCATCGGAACCCAACTTCGCGAAATGGGTCAATTAGGATTTTCGGAGAGGCCCATTTGG 141  
DB 165 ACACACAAAGCAGCAGCGCGAGTTTGGTTTCATTTTCAGCCTTTGGATCGGCCACATAC 224  
QY 142 AGGCGCGTTAATTTCAATTTATCTCAAGACGGAGCTTGTAAAGCCCGTTAAGCCGA 201  
DB 225 TATATCACTTTCGGTGAATGTTCTCGTAGACGTTTGGCCGTTAAGCCAATTAACGCTGA 284  
QY 202 ACCTCCAGCAGGATTCATTTGTTCTTCGCGACCAACCATCGTTGCTTCTGCTTC 261  
DB 285 GCCTAAAGGAATGAGTCAATTTGTCCTCGCAGCAACACCGTTGCTCCGAGTTGA 344  
QY 262 TGAGA---CGAAAGAGGAATTTTGAACAGATAGCCAGTGTATCGCAATGCTTCACC 318  
DB 345 AGAGAAAGTTGATGTGGAAGACTATGAGAAATTTGGCTGATGAGCTTCAAAATGCCCTCC 404  
QY 319 TCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACGATAGCTATTCCTTCAG 378  
DB 405 TCTTGAGATCATGGAACAAGTCCCTGGCTAAGTTTGGCAATGATATAGCTATTCCTTCAG 464  
QY 379 TGGTGTCAAGATGTTGCTTTGATTGAGTATGCGAAATTCGCGGTTCGACCTTTAGGGT 438  
DB 465 TGGGCCCCAGACGTTGCTTGTATGAGTATGACACTTGTGCTGCACTTCAGAGT 524  
QY 439 TTTCAAGTTTGGCACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTGATGGGTTGA 498  
DB 525 ATTTAGCTTGGACACAGGAGATTGAATCCGAGACATATAAAATTTTGTGATACGTTGA 584
```

```
QY 499 GAAGCATTATGGAATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTCAGGCATT 558  
DB 585 GAAACAGATATGGAATTCATATGAGTACATGTTTCTGATGCTGTGCAAGTTTCAAGCTTT 644  
QY 559 GGTGAGGAGTAAGGGGTTATTTCTTTTCTACGAGGATGGGCAACAAGAGTGTTCAGGGT 618  
DB 645 AGTTAGGAGCAAGGGGCTTTTCTTTTCTACGAAGATGGCCACCAAGAGTGTCTCCGGGT 704  
QY 619 GAGAAAGGTGAGGCTTTAAGGAGGGCCCTTAAGGTTCTCAGAGCATGGAATACTGTCGA 678  
DB 705 GAGGAAGGTGAGGCTTCCGAAAGAGCCCTTAAGGGATTACGGCATGGAATCACTGGGCA 764  
QY 679 GAGGAAGAGCAGTCACTGCTACTAGTCTCGAAATACCGGTTCTTTCAGGTTGATCCGGC 738  
DB 765 GAGGAAGATCAGTCTCTCTGCACTAGTCTGAAATTCAGTTGTCCAGTTGACCCCTGT 824  
QY 739 TTTTGAAGGGAATGGAATTTGGAAGCTTGGTGAAGTGGAAACCTGTTGCAAAATGT 798  
DB 825 TTTTGAAGGGAATGGAATGGAAGCTTGGTGAAGTGGAAATCTGTTGGCTAATGT 884  
QY 799 GAAGGCGCATGACATATGGAACCTCTTAGGACCATGAATGTGCTGTGAATTCCTTGCA 858  
DB 885 GGAAGGAAGGATATCTGGAACCTTCTCCGAGCCATGATGTGCGGTGAACATTCGA 944  
QY 859 TGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGCTGTGTTTACCTGGGCA 918  
DB 945 CTCAAGAGTTATGTTCTCAATTTGGCTGTGAGCCGTGACAGGCCAGTCTCTCCAGGGCA 1004  
QY 919 ACATGAAGGGAAGGGAGGTGGTGGGAGGATGCGCAAGCTTAAGGAATGTGCTTCTTCA 978  
DB 1005 GCACGAGAGGGAAGGAGGTGGTGGGGAAGATGCGCAAGCAAAAGGAATGTGGGTTACA 1064  
QY 979 CAAAGGAATGTAAGACAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATC 1038  
DB 1065 CAAAGGGAATATCAAG-----GAGGAACCTCTGAATTAATGGAATGGAAGTGT 1115  
QY 1039 CCATGCAAAATGGTGTATGCTACCACTGTGCTGTGACATTTTCAACAGCCCGAATGTAGTTAA 1098  
DB 1116 TAATGGAATGGAAGTGAATCTATTG---CTGATATCTTTGATACCAATATGTAACGAG 1172  
QY 1099 CTTGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGCAAGGAACCATGGCT 1158  
DB 1173 CTTGAGTAGGCTCGGATTTGAGAAATTTGCTGAAGTTGGAGGAAAGAGAGGATGAGCT 1232  
QY 1159 TGTGTGCTTTATGCAACATGCTGCTCCCTACTGCGAGCTATGAGGAACTCTTATGTTGA 1218  
DB 1233 AGTTGTACTTTTACGACCTTGGTCCCGTTTTCGCGAGCAATGGAAGGATCATATCTTGA 1292  
QY 1219 CTTAGCAGACAAGTTAGCAGGCTCAACAGGCTCAAGGTTGGAATTTAGAGCAGATGG 1278  
DB 1293 ATTGGCAGAAAGTTGGCTGGT---CTGGTGTAAAGTTGMAAGTTTAAGGCAGACGG 1349  
QY 1279 AGAA CAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTCAGATATTAT 1338  
DB 1350 TGATCAGAAAGCAATTTGCACAGCAAGAAATTCAGCTGAATAGCTTCCCTACAAATCTTTT 1409  
QY 1339 TTTCCCAAGCAATTCGCTCGGCCCAACAATAAGTATCCCTCAGAAAAAGAGATGTTGA 1398  
DB 1410 CTTCCCAAGCACTCATCTAAGC---CCATAAAATATCCATCAGAAAGAGGATGTAGA 1466  
QY 1399 TTTCTTGATGCAATTTGTAATGCTTAAGATGA 1432  
DB 1467 CTCATTGATGCTTTTGTAAATGCTTCGCGATGA 1500
```

RESULT 6

```
BT013254  
LOCUS BT013254 1835 bp mRNA linear PLN 11-MAY-2004  
DEFINITION Lycopersicon esculentum clone 134763R, mRNA sequence.  
ACCESSION BT013254  
VERSION BT013254.1  
KEYWORDS FLI CDNA.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
```

ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 1835)
 Kirksess, E.F., Wang, W. and Vazeille, A.
 Direct Submission
 Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
 LOCATION/Qualifiers
 source
 1. 1835
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 /clone="134763R"
 /tissue type="callus tissue"
 /note="TCADM70"

ORIGIN

Query Match 40.8%; Score 664.8; DB 15; Length 1835;
 Best Local Similarity 70.4%; Pred. No. 3.1e-164;
 Matches 974; Conservative 0; Mismatches 377; Indels 33; Gaps 5;

Qy 69 CTTCACCTTCCCATCATCGAAGCCAACTTCGCAAAATTCGGTCAATTAGGATTTGG 128
 Db |||||
 Qy 142 CTTTCATCTTCTTATGAAACCAAGATATCCCAATTCGGTACCTTTTCAGCCATTGG 201
 Db |||||
 Qy 129 AGAGGCCCATTTGGAGGCGCCGTTAAATTTCAATTTCTCAAGAGCGGAGCTTGTGAAGC 188
 Db |||||
 Qy 202 ATAGGCTCAACTATTGCTGCTCACTGTTTGAATTTCTCGAGGCGGTTCGCGATGAAGC 261
 Db |||||
 Qy 189 CCGTTAAAGCCGAACCTCCAGCAAGGATTCATTTGTTCTCTCGAGCAACACCATCG 248
 Db |||||
 Qy 262 CATTGTATGCTGAACCTTAAGAGGAATGATTAATAGTTCCGTCAGCAGCTACCATCGTG 321
 Db |||||
 Qy 249 TTGCTTCTGCTTCTGAGAGCAAGAGGNA---GATTTGAACAGATAGCCAGTCTCG 305
 Db |||||
 Qy 322 CTCCTGAGGTAGGAGAGAGTGTGAGGAGAGGAGCTTTTGAGAAATTTGCTTAAGGAGCTC 381
 Db |||||
 Qy 306 ACAATGCTTCACTTCTGAAATCATGATAGAGCCCTCGACAAATTCGGAACACACATAG 365
 Db |||||
 Qy 382 AATAGCTTCCCTCTTGAGGTATGACAAAGACATTTGAGAAATTTGGAGATGACATTTG 441
 Db |||||
 Qy 366 CTATTGCTTCAAGTGTCTGAGATGTGCTTTGATGATGATGCGAAATTTGACGGGTC 425
 Db |||||
 Qy 442 CTATTGCTTTCAGTGTGCTGAAGATGTTGCTTTGATAGATGACGACATTTAACTGAC 501
 Db |||||
 Qy 426 GACCTTTAGGTTTTCAGTTTGGACACTGGGAGCTGGAACCCAGAACTTATCAACTTT 485
 Db |||||
 Qy 502 GACCATACAGAGTATTGAGCTTGTATCTGGAGGTTGAACCCCGAGACATACCAATTAT 561
 Db |||||
 Qy 486 TTGATGCGGTGAGAAGCATTTATGCAATTCGATGAGTATGCTTCCCTGATGCTGTTG 545
 Db |||||
 Qy 562 TTGACAGTGGAGAGCACTATGCGCAITGCAATGATATCATGTTCCCTGATTCAGTTG 621
 Db |||||
 Qy 546 AGGTTACAGGCAATTCGTGAGGAGTAAAGGGTTATTCTCTTCTACAGGATGGGCAACCAAG 605
 Db |||||
 Qy 622 AAGTTACAGGCGTTGTTAGGACCAAGAGGGCTTTTCTCTTCTATGAGGATGGCCACCAAG 681
 Db |||||
 Qy 606 AGTGTTCAGGCTGAGAAAGTGAAGGCTTTAAGAGGCGCCCTTAAGGGTCTCAGACAT 665
 Db |||||
 Qy 682 AGTGTTCGCTGTAAAGAGGTTAGGCTTTTGAAGGAGAGCTTAAAGGGCTTACGCGCT 741
 Db |||||
 Qy 666 GGATAACTGTTGACAGAAAGACCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
 Db |||||
 Qy 742 GGATCAGGCGCAGTAAAGATCAGTCCCTGGAATCTGATGAGAAATCCCAATTTGTC 801
 Db |||||
 Qy 726 AGGTGTATCCGGCTTTTGAAGGAAATGATGTTGAAATTTGGAAGCTTGGTGAAGTGAAC 785
 Db |||||
 Qy 802 AGGTGGACCTTCTTTTGAAGGGTGTGATGCGGCTGCTGCTGCTGCTGCTGCTGCTG 861
 Db |||||
 Qy 786 CTGTTGCAATGTGAAGGCGCATGACATATGGAATCTTCTTAGGACCATGATGCTGCTG 845
 Db |||||

Db 862 CTGTGGCTTAATGTGGACGGAAAGATATTTGGAACTTCTCTGCTGCCATGAATGTGCTG 921
 Qy 846 TGAATTCCTTCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACTAGGCTG 905
 Db |||||
 Qy 922 TGAATTCATTCATTCACAGGATATGTTTCCATTGGATGCAACCTTGCACAGGCCAG 981
 Db |||||
 Qy 906 TTTTACCTGGGCAACATGAAAGGAGGAGTGTGTTGGTGGAGGATGCCAAAGCTAAGG 965
 Db |||||
 Qy 982 TTCTACCAAGGCAACACGAGAGAGGAGAGTGTGTTGGGAAAGATGCCAAGGCCAAG 1041
 Db |||||
 Qy 966 AATGTGCTTTCACAAAGGAAATGTAAGCAGCAGAAAGAGGAGGATGTTTATGGAATG 1025
 Db |||||
 Qy 1042 AGTGTGGCTTGCACAAAGGCAACATCA-----AGGATGAAACTGTAAATG 1086
 Db |||||
 Qy 1026 GGCTATCCCAATCCATGCAAAATGGTGTATGTACCACTGTGCTGCTGATATTTTCAACAGCC 1085
 Db |||||
 Qy 1087 GCGCTGCCCAACAAA-----TGGTACTGCTACCGTTGCTGATATTTTGTATACCA 1137
 Db |||||
 Qy 1086 CGAATGTAGTTAACTTCAGCAGGAGTGGAAATTTGAGAAATTTGGCAAAATTTGAGGACCGAA 1145
 Db |||||
 Qy 1138 AGGACATTTGTACCTTCAGTAAGCTGGAGTTGAGAACCTAGTAAATTTGGAAGACCGAA 1197
 Db |||||
 Qy 1146 AGGAACCATGCTTGTGCTTTATGCAACCATGTCCTTACTGCCAGGCTATGAGG 1205
 Db |||||
 Qy 1198 GAGAGCCTTGGCTGCTGTTTCTTTATGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1257
 Db |||||
 Qy 1206 AATCTTATGTTGACTTAGCAGACAAAGTTAGCAGGCTCAACAGGATGAAGGTTGGAAT 1265
 Db |||||
 Qy 1258 GATCTATGTTGAATTCGCTGAGAGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314
 Db |||||
 Qy 1266 TTAGAGCAGATGGAGAAACAGAAAGAAATTTGCAAGAGTGAACCTCAATTTGGGAAGCTTCC 1325
 Db |||||
 Qy 1315 TCAGGCGAGATGGTGACAGAGCAATTTGCAAGAGAAATTTGAGGCTTGGCAGCTTCC 1374
 Db |||||
 Qy 1326 CTAGCATATTTATTTTCCCAAGCACTTCTGCTGCGGCAACATTAAGATATCCCTCAGAAA 1385
 Db |||||
 Qy 1375 CTACAATACTCTTCTTCCCAAGCACTTTCAAAG---GCCATTAAGTACCTTTTTCAGAGA 1431
 Db |||||
 Qy 1386 AGAGAGATGTTGATTCCTTCATGCGCATTTGTAATGCTTAAAGATGAGGATATCAGGAAA 1445
 Db |||||
 Qy 1432 AGAGGAGCTAGACTCTTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1491
 Db |||||
 Qy 1446 TTTT 1449
 Db |||||
 Qy 1492 TTTT 1495

RESULT 7
 AY568717
 LOCUS
 DEFINITION Lycopersicon esculentum adenyl-1-sulfate reductase mRNA, complete cds.
 ACCESSION AY568717
 VERSION AY568717.1 GI:51457939
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 1386)
 Coker, J.S., Vian, A. and Davies, E.
 Identification, accumulation, and functional prediction of novel tomato transcripts systemically upregulated after fire damage
 Physiol. Plantarum 124 (3), 311-322 (2005)
 2 (bases 1 to 1386)
 Coker, J.S., Vian, A. and Davies, E.
 Direct Submission
 Submitted (08-MAR-2004) Botany, North Carolina State University, Gardner Hall, Campus Box 7612, Raleigh, NC 27695, USA
 LOCATION/Qualifiers
 source
 1. 1386
 /organism="Lycopersicon esculentum"

```

/mol_type="mRNA"
/db_xref="taxon:4081"
1..1386
/codon_start=1
/product="adenyl-lyl-sulfate reductase"
/protein_id="AAU03359.1"
/db_xref="GI:51457940"
/translation="MAUTFTSSSAIHGSLSSSSSYRQPKVQLGTQPLDRPOLLS
TVLNSRRSAVKPLIAPKRNDSIVPSAATIIVPEVGSVEAEDEFKELAKELQNASPL
EVMKALEKFGDDIATAFSGAEDVALIEYAHLTGHPYRVSFSLDTGRNLNPKYQLFDIV
EKHVGIRIEMFDPVSQALVRKGLFSFVDEGHQRCRVRKVRPLRALKGLRWI
TGQRKDSQPTSRPIPOVDPSFEGLDGGAGSLVKNPVANVDGDIWNFLRANVP
VNSHSQGVISIGEPCETRPVLPQCHREGRWVEDAKAKECGLHKGNIKDETVNGA
QTNGTATVATFDIKDITLSLKGPNVENVLKDLEDRPWLVLVYAPWQCFQAMGSGTV
ELAEKLAGSGVYKFKGFRADGQKAPAEELQLGSPFTILFPFKHSSKAIKYPSEKRDV
DSLILAFVNALR"
transit_peptide 1..213
/note="chloroplast transit peptide"
/evidence=not_experimental
misc_feature 346..963
/note="Region: phosphoadenosine phosphosulfate reductase
domain"
misc_feature 1048..1377
/note="Region: thioredoxin domain 2"
ORIGIN
Query Match 40.7%; Score 663.8; DB 15; Length 1386;
Best Local Similarity 70.7%; Pred. No. 5.7e-164;
Matches 967; Conservative 0; Mismatches 367; Indels 33; Gaps 5;
QY 69 CTTCCACCTCCCATCATCGGAACCAAACTCCGCAAAATGGGTCAATTAGGATTTCGG 128
DB 50 CTTTCATCTTCTTCTTATGAACCAACCAAGATATCCCAATCGGTACCTTTTCAGCATGG 109
QY 129 AGAGSCCATTTGGAGCGCGGTTAAATTTCAATTTATCTCAAGACGAGCTTGTAAAGC 188
DB 110 ATAGCCCTCAACTATTGTGCTCACTGTTTGAATCTTCGGAGGCTTCGGCAGTGAAGC 169
QY 189 CCGTTTAAACCGCAACCTCCACGCAAGGATCCATTTGTTCTCTCGCAGCAACCAACATCG 248
DB 170 CATTTGATGCTGAACCTTAAGAGGAATGATCAATAGTTCCGTCAGCAGCTACCATCGTGG 229
QY 249 TTGCTTCTGCTTCTGAGAGGAAGAGGAA---GATTTTGAACAGATACCCAGTGATCTCG 305
DB 230 CTCCTGAGGTAGGAGAGGATGTTGAGGAGAGGACTTTGAGAAATTTGCTAAGGAGCTTC 289
QY 306 ACAATGCTTCACCTCTGAAATCATGGATAGAGCCCTCGCAAAATTCGCGCAACGACATAG 365
DB 290 AAAATGCTTCCCTCTTGAGGTTATGGACAAGACACTTGAGNAATTTGGAGATGACATTG 349
QY 366 CTATTGCTTCAGTGGTGTGAAGATGTTGCTTTGATTTGAGTATCGGAAATTTGACGGGTC 425
DB 350 CTATTGCTTCAGTGGTGTGAAGATGTTGCTTTGATAGTATGACACATTTAACTGGAC 409
QY 426 GACCTTTAGGGTTTTCAGTTTGGACACTGGAGACTGAACCCAGAACTTATCAACTTT 485
DB 410 GACCATACAGAGTATTCAGCCTTGATATCTGGGAGGTGAACCCGGAGACATACCAATTAT 469
QY 486 TTGATCGGTTTGAGAAGCAATTATGAATTCGCATTGAGTATGATTTTCTCTGATGCTGTG 545
DB 470 TTGACACAGTGGAGAAGCACTATGGCATTCGCATTGATATGATTTCTCTGATTCAGTTG 529
QY 546 AGGTTTCAGGCATTGTGGAGAGTAAGGGGTTATTTCTTTTTCAGGAGATGGCCACCAAG 605
DB 530 AAGTTTCAGGCGTTGTAGGACCAAGGGCTTTTCTTTCTTATGAGGATGGCCACCAAG 589
QY 606 AGTGTTCAGGCTGAGAAGGTGAGCCCTTAAAGGAGCCCTTAAGGCTCTCAGGCGAT 665
DB 590 AGTGTTCGCGTTAAGAGAGGTATAGCCCTTTGAGGAGAGCTCTAAGGGCTTACGCGCT 649
QY 666 GGATAACTGGTCAGAGGAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGTTGTTC 725
DB 650 GGATCACAGGCCGTTAAGATCAGTCCCTTGGAACTCGATCAGAAATCCCCATTTGTC 709

```

```

QY 726 AGCTTGATCGGCTTTTGAGGGAATGATGTTGGAATTTGGAAGCTTGGTGAAGTGGAAACC 785
DB 710 AGTGGGACCTTCTTTTGAGGGTTGATGCGGTTGCTGTAGCTTGGTGAAGTGGAAACC 769
QY 786 CTGTTGCAAAATGTGAAGGGCCATGACATATGGAACTTCTCTTTCAGGACCATGATGTCCTG 845
DB 770 CTGTGGCTAATGTGGACGGAAGATATTTGGAACTTCTTCTGCGTGCATGATGTCCTG 829
QY 846 TGAATTCCTTGATGCAAAAAGGATATGTTTCCATTTGGTGTGAGCCCTGCACTAGGCTG 905
DB 830 TGAATCTATTGCAATTCACAGGATATGATTCATTTGGATGCGAACTTTGCAACAAGGCCAG 889
QY 906 TTTTACCTCGGCAACATGAAGGCAAGGAGGTGCTGTGTGGAGATGCCAAAAGCTAAGG 965
DB 890 TTTTACCAGGCAACACGAGAGAGAGGGAAGATGTTGTGTGGAGATGCCAAGGCCAAG 949
QY 966 AATGTGCTCTTTCACAAAGGAAATGTAAGCAGCAGAAAGAGGAGATGTTTAAATGGAATG 1025
DB 950 AGTGTGGCTTGCAACAAGGCCAATCA-----AGGATGAAACTGTAAATG 994
QY 1026 GGCTATCCCAATCCCATGCAAAATGGTGATGCTACCATGTCCTGACATTTTCAACAGCC 1085
DB 995 GCGCTGCCCAACAAA-----TGGTACTGCTACCGTTGCTGATATTTTGTATACCA 1045
QY 1086 CGAATGATGTTAACTTTCAGCAGGACTGGAATTTGGAATTTGCAAAATTTGAGGAGCCGAA 1145
DB 1046 AGGACATTTGTACCTTGAGTAAGCTGGAGTTGAGAACCTAGTAAATTTGAGAGCCGAA 1105
QY 1146 AGGAACCATGCTTGTGCTTTTATGCACCATGCTGCCCCCTACTGCGAGGCTATGGAGG 1205
DB 1106 GAGAGCCTTGGCTGTTGTTCTTTATGCACCTTGGTGCCAATTTTGCAGGCAATGGAAG 1165
QY 1206 AATCTTATGTTGACTTAGCAGACAAAGTTAGCAGGCTCAACAGGATGAAGGTTGGAAT 1265
DB 1166 GATCTATGTTGAAATTTGGCTGAGAAGTTGGCTGTT---CTGGTGTGAAGATGAGGNAAT 1222
QY 1266 TTAGAGCAGATGGAGACAGAAAGAAATTTGCAAGAGTGAAGTCAATTTGGGAAGCTTCC 1325
DB 1223 TCAGGGCAGATGGTGACCAAGAACATTTGACAGAAGAAATTTGACAGCTTGGCAGCTTCC 1282
QY 1326 CTACGATATTTATTTTCCCAAGCAATTCGCTCTCGGCCAACAAATAAGATATCCCTCAGAAA 1385
DB 1283 CTACATATCTTCTTCCCAAGCACTCTTCAAAG---GCCATTAACTACCTTTCAGAGA 1339
QY 1386 AGAGAGATGTTGATTCCTTGATGTCATTTGTAATGCTTAAATGCTTAAAGTGA 1432
DB 1340 AGAGGAGCTAGACTCCTTGTCTGCTGCTTTTGTGAATGCTCTCAGATGA 1386

RESULT 8
SCJ001208          Brassica juncea mRNA 1608 bp mRNA linear PLN 21-MAY-1999
LOCUS              Brassica juncea mRNA for APS reductase.
DEFINITION         AJ001208
ACCESSION          AJ001208.1 GI:4033576
VERSION            APS reductase.
KEYWORDS            Brassica juncea
SOURCE              Brassica juncea
ORGANISM            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                    rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1
REFERENCE           Heiss,S., Schafer,H.J., Haag-Kerwer,A. and Rausch,T.
AUTHORS             Cloning sulfur assimilation genes of Brassica juncea L.: cadmium
TITLE               Differentially affects the expression of a putative low-affinity
                    sulfate transporter and isoforms of ATP sulfurylase and APS
                    reductase
JOURNAL             Plant Mol. Biol. 39 (4), 847-857 (1999)
PUBMED              10350097
REFERENCE           2 (bases 1 to 1608)
AUTHORS             Heiss,S.
TITLE               Direct Submission

```

JOURNAL Submitted (15-JAN-1998) Heiss S., Ruprecht-Karls-Universitaet
Heidelberg, Botanical Institute, Im Neuenheimer Feld 360, D-69120,
GERMANY

FEATURES
source Location/Qualifiers
1..1608
/organism="Brassica juncea"
/mol_type="mRNA"
/cultivar="Vitasso"
/db_xref="taxon:3707"
/tissue_type="root"
1..1608
/gene="apsr8"
30..1424
/gene="apsr8"
/note="putative mitochondrial or plastidic transit
peptide"
/codon_start=1
/product="APS reductase"
/protein_id="CAA04611.1"
/db_xref="GI:4033577"
/db_xref="GOA:Q92P22"
/db_xref="UniProt/TcEMBL:Q92P22"
/translation="MALAINVSSSSSAISTSPSSSELKVASPRIGSLRLSDRVNVS
SVLSLGGKSSVKPLNVQSIARKEVSFGSMVASEVTERKQDVVEDEFEELAKSLTASP
LEIMDKALETFGNIDIAIFAGAEVDALIEYAHLTGRFVRVSLDTGRNLNPEYRLFT
VEKHYGIRIEMFPDAVEQALVRNKLGFSEYEDHOECRIRKVRPLRLALKGLRAW
ITGQRKDQSPETSEIPIVQVDVPVPEGLDGGAGSLVKWNPVANVGNVDFLRTMDV
PVNTLHAAGVSVISGCEPCTRAVLPGOHEREGRWWEAKAKEGCLHKGNIKESSGNA
TANVNGSVTVADI FKSENVSLSROGIENLMLENRKEAWIVLVYAPWPCFPQAMBA
SPLEADKLGGSGVKVAKPRADGQKEFAKRELQGLSFPFILVFPKNSRPIKPYSEK
RDVSLTSLFLNLVR"

transit_peptide 30..194
/gene="apsr8"
/note="putative mitochondrial or plastidic transit
peptide"

mat_peptide 195..1421
/gene="apsr8"
/product="APS reductase"
/note="putative mitochondrial or plastidic transit
peptide"

ORIGIN

Query Match 40.7%; Score 663.6; DB 15; Length 1608;
Best Local Similarity 68.7%; Pred. No. 6.4e-164;
Matches 985; Conservative 0; Mismatches 419; Indels 30; Gaps 4;
QY 32 ATGCGCCCTCGCTTCACTTCTCAATTTCCGCCACCACTTCCACCTTCCCATCATCGAA 91
DB 42 ATCAACGTTTCTCATCTTCTTCTTCTCGCGATCTCAACCTCTAGCTTCCCTTCTCAGAG 101
QY 92 CCCAAACT-----TCCGCAAAATGGGTCAATTAGGATTTCCGAGAGGCCCATTTGGAGGC 145
DB 102 CTCAAGTTTCAGCTCCACGGATCGGTTCGTGGGTATTCGGATCGTGTCATGTCTCA 161
QY 146 GCGGTAAATTCAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACGCCGACCT 205
DB 162 TCGGTGTCTCTAGCTCTATCGGGAACGATCTCTCGGTGAAGCTCTGAATGTGCACTCA 221
QY 206 CCAGCGAAGATTCATTTGCTCTCTCGCAACCAACCATCGTTGCTTCTGCTTCTGAG 265
DB 222 ATTGCAAGAGAGTGGTTGGTTCTATGGTGGCTTCTGAGGTTACAGAAACAAGATGTG 281
QY 266 ACAGAAAGAGGAATTTGAACAGATAGCCAGTGATCTCGACAAATGCTTCACCTCTTTGAA 325
DB 282 GTGGAGTGAAGACTTTGAGGAGCTAGCAAGAGTCTAGAGCTGCTTCTCTCTTTGAG 341
QY 326 ATCATGGATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGCTTTCAGTGGTGTCT 385
DB 342 ATTATGACAAAGCTCTTGAGACATTCGGAACGACATCGCAATCGCTTTTAGTGGAGCA 401
QY 386 GAAGATGTTGCTTGTAGTATGCGAAATTCGAGGGTCGAGCCCTTTAGGGTTTTCAGT 445
DB 402 GAAGACGTTGCTCTCATCGAGTATGCTCATCTTACTGGAAGACCCCTTCAGGGTTTTCAGT 461

RESULT 9
AP004494/c
LOCUS

AP004494 103129 bp DNA linear PLN 22-JUL-2003

QY 446 TTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGGGTTGAGAAGCAT 505
DB 462 TTGGACACAGGAGATTGAATCCGGAACATACAGACTCTTTGTATACCGTGGAGAGCAC 521
QY 506 TATGGAATTCGCATTGAGTACATGTTCCCTGTAGTCTGTTGAGGTTCAAGGCATTTGGTAGG 565
DB 522 TATGGTATTGGAATTTGAGTATATGTTTCCCGATGCTGTTGAGGTTCAAGCTTTGGTTAGG 581
QY 566 AGTAAGGGGTTATTCTCTTTACGAGATGGGGCACCAGAGTGTTCAGAGGTGAGAAAG 625
DB 582 AACAGGGTTTGTCTCTTTACGAAGATGGTACCAGAGTGTTCGCGATCAGAAAG 641
QY 626 GTGAGGCTTTTAAAGAGGGGCCCTTTAAGGGTCTCAGAGCATGGATAACTGTGTCAAGGAAA 685
DB 642 GTGAGACTTTTGAAGGCGTGGTTGAAGGGGTACGCGCTGGATCACTGGACAAAGGAAA 701
QY 686 GACCACTCAGCTGTGTAAGTCTGTAATAACCGGTTTGTTCAGGTTGATTCGGCTTTTGTAG 745
DB 702 GATCAATCACAGGAGCAAGGCTCTGAGATTCCTGTTCTTCAAGTTGATCCAGTCTTTGAA 761
QY 746 GGAATGATGCTGGAATTTGGAAGCTTGGTGAAGTGGAAACCTGTTCCAAATGTGAAGGC 805
DB 762 GGGTTAGATGGTGGAGCTGGTAGTTTGGTGAAGTGGAAATCCAGTTGCGAAACGTCGAGGA 821
QY 806 CATGACATATGGAACCTTCTTAGGACCATGAATGTGCTGTGAATTCCTTGCATGCAAAA 865
DB 822 AACGATGTTTGAACCTTCTTAGGACCATGAATGTGCTGTGAACACCTTCAAGCTTGG 881
QY 866 GGAATGTTTTCATTTGGTGTGAGCCCTGCACTAGGCTCTTTTAACTCTGGGCAACATGAA 925
DB 882 GGGTATGTTTCAATCGGATCGAGCGGTGACGAGAGCGGTTTTCCTGCTCAGCAGAG 941
QY 926 AGGGAAGGAGGTGGTGGGAGGATGCCAAGCTTAAGNAATGTGCTCTTCCAAAGGA 985
DB 942 AGAAGAGGAGATGGTGGGAGATGCTTAAGGCTTAAGAGTGTGGACTTTCACAAAGGG 1001
QY 986 AATGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCATGCA 1045
DB 1002 AATATCAGGAG-----ASTAGCAACGGAACGCTACCGCTAATGTC 1043
QY 1046 AATGGTATGCTACCACTGCTGCTGACATTTTCAACAGCCGCAATGTAGTTAACTTGAGC 1105
DB 1044 AATGGCACATCTCTCCACGGTTGCTGATATCTTCAAAAGCGAGAATGTGTGAGCTTGAGC 1103
QY 1106 AGGACTGGAAATTCAGAAATTTGGCAAAATTCGAGGACCGGAAGGAAACCATGGCTGTTGTG 1165
DB 1104 AGCAAGGGGATCAGAAATCTGATGAAGCTGGGAATCGTAAAGAGGCTTGGATCGTTGTG 1163
QY 1166 CTTTATGCACTATGGTCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCA 1225
DB 1164 CTTTACGACCTTGGTCCCGTTTGTCAAGCGATGGAGCTCTTTTGTATGAGTTGGCG 1223
QY 1226 GACAAATTTAGCAGGGTCAACAGGGATGAAGTTGGAAAAATTTAGAGCAGATGGAGAACAG 1285
DB 1224 GATAAGTTGGGTGGG---AGTGGCGTGAAGGTGCCAAGTTTAGAGCTGATGGTGAACAA 1280
QY 1286 AAGAATTTGCAAGAGTGAACCTGCAATTTGGGAGCTTCCCTACGATATTATTTTTCCTCA 1345
DB 1281 AAGGATTTGCTTAAAGAGGAGCTGCAAGCTTGGGAGCTTCCCAACAATATCTGTTGTTCCCA 1340
QY 1346 AAGCATTTGCTCGGCCCAACAATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTG 1405
DB 1341 AAGAACTCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGATGTTGATTCCTTG 1397
QY 1406 ATGGCAATTTGTAATGCTTTAAGATGAGGATATCAGGAAATTTTTCCTGTTTTT 1459
DB 1398 ACATCGTTCTTGAACCTTGTTCGGTGAATAATAGCAAAAATCCGTCGACTTTTT 1451

Qy	854	TTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACATAGCCCTGTTTTACCT	913
Db	42869	TTGCATTTCAAAAGGATACATCTCCAATTGGCTGCGAGCCGTGCACGAGGGCTGTTTTACCC	42810
Qy	914	GGGCAACATGAAAGGGAAGGGAGGTGGTGGTGGGAGGATGCAAAAGCTTAAGGAATGTGGT	973
Db	42809	GGGCAACATGAAACGGGAAGGTAGTGGTGGTGGGAGGATGCAAAAGCTTAAGGAGTGTGGT	42750
Qy	974	CTTTCACAAAGGAAATGTTAAAGCAGCAGAAAGAGGAGGATGTTAATTGGAAATGGGCTATCC	1033
Db	42749	CTTTCACAAAGGAAATGTGAAGCAGGATCGGAGGCGGAGCTTAACGGCAATGGTGTGGCC	42690
Qy	1034	CAATCCCATGCAAAATGGTGATGCTACCACTGTGCTGTGACATTTTCAACAGCCCGCAATGTA	1093
Db	42689	AATACAAATGGCA-----CTGCCACTGTGCGCCGACATCTTCAACACCCAGAAATGTG	42639
Qy	1094	GTTTAACCTCAGCAGGAGCTGGAAATCAGAAATTTGGCAAAATTTGGAGGACCCGAAAGAACCA	1153
Db	42638	GTCAAGCTTGAGCAGGAGCTGGAAATTCAGAAATTTGACAAAGCTTGGAGAACCGGAAGAACCA	42579
Qy	1154	TGGCTTTGTTGTGCTTTATTCACCACTGGTGCCTTACTGC-----	1192
Db	42578	TGGTTGGTTGTGCTTTATTCACCACTGGTGCCTTACTGCCAGGTAACATAATTTTCTGTC	42519
Qy	1193	-----	1192
Db	42518	AGAAATTTGTAGCTGAGCTCATTAACATATGTTTACCTTGGGTGGAATGTTTGATCCCTA	42459
Qy	1193	-----CAGGCTATGAGGAACTCTTATGTTGACTTAGCAGA	1227
Db	42458	ATGAATTTCTATTCGAAATTCGAAACAGGCTATGAGGAACTCTTATGTTGATCTGGCTGA	42399
Qy	1228	CAAGTTTAGCAGGCTCAACAGGAGTGAAGGTTGGAAAAATTTAGACAGATGGAGAACAGAA	1287
Db	42398	CAAGTTTAGCAGAGT---CAGGGGTGAAAGTTGGAHAATTTAGACGGATGGGAGCAGAA	42342
Qy	1288	AGAATTTGCAAAAGAGTGAACTGCAATTTGGGAAGCTTCCCTACGATATTATTTTCCCAAA	1347
Db	42341	AGAATTTGCAAAAGCGGGAACCTGGGGTTGGGAGCTTCCCAACAAATAATGTTTTTCCCTTAA	42282
Qy	1348	GCATTCGTCTCGGCCAACAAATTAAGTATCCCTCAGAAAGAGAGATGTTGATCCCTCGAT	1407
Db	42281	GCATTCATCTAGGC---CAATAAAGTATCCCTCAGAAACAGAGATGTTGATTCCTTGAT	42225
Qy	1408	GGCAATTTGTAATGCTTAAAGTATGAGGATATCAGGAAATTTTCTTCGTTTTTGGGTGCA	1467
Db	42224	GGCGTTTGCAATGCCCTTAAAGTAGGGTTATCAAGAGTTCTT---TGTTTTAGGTGCC	42168
Qy	1468	ATTCCACTTTTGACTATACGTACAGCGGGTTCTTCTTTA	1506
Db	42167	ATTCAATTTGTAGTGACCTTCAATGTTCTCTTGCTCA	42129

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Arabidopsis ORF clones
 Unpublished
 2 (bases 1 to 1377)
 Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayaishizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Direct Submission
 Submitted (11-MAY-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'; Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayaishizaki, Y. and Shinozaki, K.)
 The Salk, Stanford, PSEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.
 Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.
 Location/Qualifiers
 1. .1377
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="U13952"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51"
 CDS
 1. .1377
 /note="PRH26 protein"
 /codon_start=1
 /product="AT4g21990/FLN20_90"
 /protein_id="AAW65364.1"
 /db_xref="GI:21655305"
 /translations="MALAINVSSSSAIISSSSPSSDLKVTIKIGSLRLNRTNVSAA
 SLSLGRSSVKALNVQSIKESIVASEVTEKLDVVEDEPELAKRLNLSAPLEIMD
 KALEKGNDAIAFSAEDVALLEYAHLGRPYRFSLDTRGLNPTETLFDVEKHY
 GRIQYFNPDAVQALVRNKLGLFSYEDHQECIRKVRPLRLAKGLRWITQOR
 KDQSPGTSIEIPVQDVPFEGDGLSVLKNVPANVEGNDVNFRLTMDVPVNTL
 HAQYVSGICEPTNVLPLQHEGRGKRWMDAKCEKGLHKNIKENTNGNATNVN
 GTASVADI FNSENVNLRSQIENLMLKLNREKATWLVYAPWCPFOAMEASFDELA
 DKLGSQVKVAKFRADGDKDFAKELQLGSFTLILVFPKNSRPKIPSEKRDVSL
 TSFLNLR"
 ORIGIN
 Query Match 40.3%; Score 657.2; DB 15; Length 1377;
 Best Local Similarity 69.5%; Pred. No. 3.2e-162;
 Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;
 31 AATGGCCCTCGCTTTCATTCTTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGA 90
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 12 AATCAAGGTTCTTCT 71
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 91 ACCCAAACTCCGCAATTTGGGTCAATAGGATTTCCGAGAGCCCATTTGGAGGCCCGT 150
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 72 TCTCAAAGTAACAAAATCGGATCATTTGAGGTTATTGAATCGTACCAATGTAATTTGGCGG 131
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 151 TAATTTCAATTTATCTCAAGAGCGGAGCTTTGGTAAAGCCGCTTAAACCCGCAACCTCCACG 210
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 132 TTCTCTGAGTTTGTCCGGGAAGAGATCTCTCCGTGAAGCTCTTAATGTGCAATCAATTAC 191
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 211 CAAGGATTCATTTGTTCTCTCGAGCAACAACCATCGTTGCTTCTGCTCTGAGAGAA 270
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 192 AAAGGATTCATTTGTTCTCTGAGGTTACAGAGAAGCTAG-----ATGTGGTGA 242
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 271 AGAGGAAGATTTTGAACAGATAGCAGTATCTCGAAGTCTTCACTCTCTGGAATCAT 330
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 243 AGTTGAAGATTTTGAAGAACTAGCAAGAGATTAGAGAATGCTTCTCTCTTGAGATCAT 302
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 331 GGATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGCTTCACTAGTGTGCTGAAGA 390
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 303 GGATAAGCTCTTGGAGAAGTTTGGGAATGACATTCCTATTGCTTCTTGTAGTGAGCTGAAGA 362
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 391 TGTGCTTGAATTCGGAATTTGAGGGTTCGACCTTTAGGGTTTTTCAGTTTGGGA 450
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 363 TGTGCTCTGATTCGATGCTCATTTAACTGGAAGACCTTATAGGCTATTTCAGTTTGA 422
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 451 CACTGGGAGACTGAACCCGAAACTTATCAACTTTTGTATGCGGTTGAGAGCAATTATGG 510
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 423 TACAGGAGATTGAATCCAGAAACATATAGACTCTTCGATACCCTGGAGAAACATTACGG 482
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 511 AATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTGCTGAGGAGTAA 570
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 483 TATTCGATTGAGTATATGTTCTCTGATGCTGTTGAGGTTCAAGCTTTAGTTAGGAACAA 542
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 571 GGGTTATTCTCTTCTACAGAGATGGGCAACAGAGTGTTCAGGGTGAAGAGGTGAG 630
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 543 GGGTTGTTCTCTTCTATGAAGATGGTCAACAGGAGTGTTCGCTATTATGAAGAGGTGAG 602
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 631 GCCTTTAAGGAGGCGCTTAAGGCTCTCAGAGCATGATAACTGCTCAGAGGAAGACCA 690
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 603 ACCTTTGAGCGGTGCGTTGAAGGGTTTACCTGCTTGGATCACCTGGTCAGAGGAAGATCA 662
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 691 GTCACCTGTGTAGGTCTGAAATACCGGTGTTTTCAGGTTGATCCGGCTTTTTCAGGGAAT 750
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 663 ATCACCGGAGACAAGATCGAGATTCCTGTTGTTTTCAGGTTGATCCAGTGTTCAGGATTT 722
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 751 GGATGTGGAATTTGGAAGCTTGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCATGA 810
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 723 AGATGTGGAGTTGGTAGTTTGGTGAAGTGAATCCGGTTGCTTAATGTTCAAGGGAATGA 782
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 811 CATATGGAATTCCTTTAGGACCATGATGTCCTGTGAATTCCTTGCATCAAAAGGATA 870
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 783 TGTTTGAATTTCTTGAGGACGATGATGTTCTCTGCAACACACGCTTCACCGCTGCGGGTTA 842
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 871 TGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGCCCTGTTTTTACCTGGGCAACATGAAAGGA 930
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 843 TGTTTCAATCGGATGTGAGCCATGCACGAGAGCGGTTTTTCCCGGTCAGCATGAGAGA 902
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 931 AGGAGGTGTGTGGGAGGATGCCAAAGCTAAGGAATGTGGTCTTTCACAAAGAAATGT 990
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 903 AGGAGATGTGTGGGAAAGATGCTAAGGCTAAAGAGTGTGGGCTTTTCAACAAGGGAATAT 962
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 991 AAAGCAGCAGAAAGAGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCCAATGG 1050
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 963 CAAGGAGAATACCAATGGAAACCGAACCCGCTAATG-----TCAA 1001
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGGAATGTAGTTAACTTTGAGCAGGAC 1110
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1002 CGGACAGCTCTCAGTCGCGGATATCTTCAATAGCAGAGATGTTGTGNACTTTGAGCAGCA 1061
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1111 TGAATTTGAATTTGGCAAAATTTGAGGACCGAAAGAACCATGGCTTGTGTGCTTTA 1170
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1062 AGGATTTGAATTTGATGAAGTTAGAGAACCGGAAAGAGGCTTGGATCGTTGTGCTTTA 1121
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1171 TGCACCATGCTGCCCCCTACTGCGAGGCTATGGAGGATCTTATGTTGACTTTAGCAGACAA 1230
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1122 TGCACCGGTGTCGCCGCTTTTGTTCAGGCAATGGAAGCTTCATTCGATGAATTTGGCGGATA 1181
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 783 TGTGTAATTTCTTGAGGACGATGGATGTTCTGTCAACACGCTTCACGCTCGCGGTTA 842
 Qy 871 TGTTCATTTGGTGTGAGCCCTGCACTAGGACCTGTTTACCTGGGCAACATGAAAGGGA 930
 Db 843 TGTTCATCGATGTCAGCCATGACGAGAGCGGTTTTCGCGGTGAGATGAGAGA 902
 Qy 931 AGGAGTGTGTGGGAGAGTGCACAGCTTAAGGATGTGTTTTCACAAAGGAATGT 990
 Db 903 AGGAGATGTGTGGGAGATGCTAAGGCTTAAGAGTGTGGCTTCACAAAGGAATAT 962
 Qy 991 AAAGCAGCAGAAAGAGAGAGATGTTAATGGAATGGGCTATCCAAATCCCATGCAATGG 1050
 Db 963 CAAGGAGATACCAATGGAACGCAACGCGTAATG-----TCAA 1001
 Qy 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTTGAGCAGGAC 1110
 Db 1002 CGGACAGCCTCAGTCGCGGATATCTTCAATAGCGAGATGTTGTGAACCTTGACGAGCA 1061
 Qy 1111 TGGATTTGAGAAATTTGGCAAAATTTGAGAGCCGGAAGAACCAATGCTGTGTGCTTTA 1170
 Db 1062 AGGATTTGAGAAATTTGATGAATTTAGAGAACCGGAAGAGGCTTGGATCGTGTGCTTTA 1121
 Qy 1171 TGCACCATGTGCCCCCTACTGCCAGGCTATGAGGAATCTTATGTTCACTTAGCAGACAA 1230
 Db 1122 TGACCGGTGTGCCCCGTTTGTGAGGCAATGGAAGCTTCAATCGATGAATTTGGCGGATAA 1181
 Qy 1231 GTTAGCAGGCTCAACAGGATGAAGGTTTGGAAATTTAGAGCAGATGAGAGAACAGAAAGA 1290
 Db 1182 GCTGGGAGG--AAGTGTGTGAAGTTGCTAAGTTTAGAGCTAGTGTGACCAAGAGA 1238
 Qy 1291 ATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTTACGATATATTTTCCCAAGCA 1350
 Db 1239 CTTTGGCAAAAGAGGTTGCAAGCTTGGAGCTTTCACAAATATCTTGTGTCCCAAGAA 1298
 Qy 1351 TTGCTCTCGCCACATTAAGTATCTCCTCAGAAAGAGAGATGTTGATTCCTTGATGGC 1410
 Db 1299 CTCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGATGTTGATTCCTTGATC 1355
 Qy 1411 ATTTGTAAT 1420
 Db 1356 TTTCTGAAT 1365
 RESULT 12
 AY062665
 LOCUS
 DEFINITION
 Arabidopsis thaliana PRH26 protein (At4g21990; Fln20.90) mRNA, complete cds.
 ACCESSION
 AY062665
 VERSION
 AY062665.1
 KEYWORDS
 FLI CDNA.
 SOURCE
 Arabidopsis thaliana (thale cress)
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1654)
 Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Bath, J., Carninci, P., Chen, H., Kawai, J., Kim, C., Lin, J., Liu, J., Li, S.X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.
 Direct Submission
 Submitted (14-NOV-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 e-mail for correspondence: arab@sequence.stanford.edu
 COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : "RIKEN Arabidopsis Full-length cDNA") : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Palm, C.J., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Bath, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, J., Li, S.X., Pham, P.K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. 1654
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="IV"
 /clone="RAFL08-16-K20"
 /ecotype="Columbia"
 /notes="This clone is in pBluescript"

gene

1. 1654
 /genes="At4g21990; Fln20.90"
 77. 1453
 /genes="At4g21990; Fln20.90"
 /codon_start=1
 /product="PRH26 protein"
 /protein_id="AAL32743.1"
 /db_xref="GI:17065178"

CDS

/translation="MALAINVSSSSSAISSSPSSDLKVTIKIGSLRLNRNVSAA
 SLISGKRSVKAIVNQSIKESIVASEVTEKLDVVEDEFEELAKLELASPLEIKH
 KALEKFNDAITAFSGAENDVALTEYAHLTGRPVFSLDTRLNPETRYLFDVTEKH
 GRIEYMFDAVEAQRNKLGLFSEVDGHQSCRRKVRPLRLALKGLRAWTTGOR
 KQSPCTREIPLVQVDPVEGLDGGVSLVKNPVANVEGNDWNPFLRMDVPNTL
 HAQYVSGCEPCTRAVLPGHREGRMEDAKAKEGLHKGINKENTNGNATANN
 GTASVADIFNSENVNLRSQIENLMKLENKEAWIVLYAPMCPFCQAMBSFDELA
 DKLGSGGVKAFKPRADGDDQDFAKELQLGSFPTILVFPKNSRPIKYPSEKRDVSL
 TSLNLR"

ORIGIN

Query Match 40.3%; Score 657.2; DB 15; Length 1654;
 Best Local Similarity 69.5%; Pred. No. 3,2e-162;
 Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;

Qy 31 AATGGCCCTCGCTTTCACCTTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGA 90
 Db 88 AATCAACGTTTCTTCATCTTCTTCTTCTGCGATCTCAAGCTCTAGCTTCCCTTCTTCA 147
 Qy 91 ACCCAAACTTCCGCAAAATTTGGGTCAATTTAGGATTTTCGGAGAGGCCCAATTCGAGCGCGT 150
 Db 148 TCTCAAGTAACAAATAATCGGATCATTTAGGTTATTGAATCGTACCAATGCTCTGCGGC 207
 Qy 151 TAATTTCAATTTATCTCAAGAGCGAGCTTGGTAAAGCCCGTTAAAGCCGGAACCTCCACG 210
 Db 208 TTCTCTGAGTTTGTCCGGGAAGAGATCTCCGTTGAAAGCTCTTAATGTGCAATCAATTAC 267
 Qy 211 CAGGATTCATTTGTTCTCTCCGACCAACCACTGTTGCTTCTCTCTGAGACGAA 270
 Db 268 AAGGAATTCATTTGTTGCTTCTGAGGTTTACAGAGAAGCTAG-----ATGTGGTGA 318
 Qy 271 AGAGGAAGATTTTGAACAGATAGCCAGTGTATCTCGACAATGCTTCACTCTCTGAAATCAT 330
 Db 319 AGTTGAAGACTTTGAGGAAGTACGAAAGAGATTAGAGATGCTTCTCTCTTGAATCAT 378
 Qy 331 GGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTCAAGTGGTGTGGAAGA 390
 Db 379 CGATAAAGCTCTTGAAGAAGTTTGGGAATGACATTGCTATTGCTTGTAGTGGAGCTGAAGA 438
 Qy 391 TGTGCTTTCATTTGAGTATGCGAAATTTGACGGTTCGACCTTTAGGGTTTTCAGTTTGA 450
 Db 439 TGTGCTTTCATTTGAGTATGCTCAATTTAACTGGAAGACCTTATAGGGTATTTCAGTTTGA 498

```
QY 451 CACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGATCGCGTTGAGAACGATTATGG 510
Db 499 TACAGGAGATTTGAATCCAGAAACATATAGACTCTTCGATACCTGGAGAAACATTACGG 558
QY 511 AATTTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCAATTTGGTGAGGATAA 570
Db 559 TATTCGATTCAGTATATGTTCTCTGATGCTGTTGAGGTTTCAGGTTCAAGCTTTAGTTAGAACAA 618
QY 571 GGGGTTATCTCTTCTACAGAGATGGGCCAACAGAGTGTTCAGGGTGAGAAAGGTGAG 630
Db 619 GGGTTTCTCTCTTCTATGAAGATGCTCACCAGAGTGTTCGCGTTATTTAGAAAGGTGAG 678
QY 631 GCCTTTAAGGAGGCCCTTAAGGCTCTCAGAGATGATGATCTGTCAGAGGAAGAACCA 690
Db 679 ACCTTTGAGGCGTGGTTGAAGGGTTTACGTGCTTGGATCTGTCAGAGGAAGATCA 738
QY 691 GTCACTCTGTTACTAGGCTCTGAATACCGGTTGTTTTCAGGTTGATCCCGGTTTTCAGGGAAT 750
Db 739 ATCAACCCGGGACAGATCGGAGATCTCTGTTGTTTCAGGTTGATCCAGTGTTCAGGAT 798
QY 751 GGATGGTGAATTCGAAAGCTTGGTGAAGTGAACCCCTGTTTCGCAATGTGAAGGCCATGA 810
Db 799 AGATGGTGGAGTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 858
QY 811 CATATGGAATCTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATGCAAAAGGATA 870
Db 859 TGTTCGAAATTTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATGCAAAAGGATA 918
QY 871 TGTTCGAAATTTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATGCAAAAGGATA 930
Db 919 TGTTCGAAATTTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATGCAAAAGGATA 978
QY 931 AGGAGTGTGTGGGAGGATGCAAGAGCTTAAGGATGTCCTTCAAAAGGATGTT 990
Db 979 AGGAGTGTGTGGGAGGATGCAAGAGCTTAAGGATGTCCTTCAAAAGGATGTT 1038
QY 991 AAAGCAGCAGAAAGGAGGATGTTAAGGATGTCCTTCAAAAGGATGTT 1050
Db 1039 CAAGGAGAAATCAACCAATGGAACCGAACCGCTAATG-----TCAA 1077
QY 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGCAATGTAGTTTACCTTGAGCAGGAC 1110
Db 1078 CGGACAGCTCTAGTCCGCGATATCTTCAATAGCGAGATGTTGTGAATTTGAGCAGGCA 1137
QY 1111 TGAATTCAGAAATTTGCAAAATTTGAGAGGACCGAAAGCAATGCTGTTGTGCTTTA 1170
Db 1138 AGGATTCAGAAATTTGATGAAGTTAGAGAACCGGAAAGAGGCTTGGATCGTTGTGCTTTA 1197
QY 1171 TGCACCATGTTGCCCTTACTGCCGCTATGAGGAGATCTTATGTTGACTTTAGCAGACAA 1230
Db 1198 TGCACCGTGTGCTGCTGTTTGTGAGGCAATGGAAGCTTCAATCGATGAATTTGGCGGATA 1257
QY 1231 GTTAGCAGGTCACACAGGATGAAGTTTGGAAATTTTAGAGCAGATGAGAGAACAGAAAGA 1290
Db 1258 CTGGGAGG---AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGTCAGCAGAGAGA 1314
QY 1291 ATTTCAGAAAGTGAATTCGAAATTTGGGAGGCTTCCCTACGATATATATTTTCCCAAAGCA 1350
Db 1315 CTTTTCGCAAAAGGAGTTGCGAGCTTGGAGGCTTTCACCAATACCTTGTGTTTCCCAAAGAA 1374
QY 1351 TTCGCTCGGCCCAACAAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTCATGCGC 1410
Db 1375 CTCCTTCAAG---ACCAATCAAGTATCCATCAGAGAAAGAGAGATGTTGATTCCTTCATC 1431
QY 1411 ATTTTGAAT 1420
Db 1432 TTTCTTGAAT 1441
```

RESULT 13
AY087591
LOCUS Arabidopsis thaliana clone 36866 mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 36866 mRNA, complete sequence.

ACCESSION AY087591.1
VERSION GI:21406329
KEYWORDS FULL CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1702)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troupkan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
PUBMED 12093376
REFERENCE 2 (bases 1 to 1702)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1702)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES
source

1. 1702
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clones="36866"
78. 1454
/codon_start=1
/product="PRH26 protein"
/protein_id="AAM5133.1"
/db_xref="GI:2159184"
/translation="MALAINVSSSSSAISSSPSSDLKVTKIGSLRLNRTNVSAA
SLSGKRSVSKALNVQISIVASEVTEKLIDVEDEFEELAKRLNLSPLIMD
KALRGNDIAIAFGSDVALIEVHLTPRPYRVSIDTGLNPEYRLFDTVEKHY
GIRLEYMPDAVEQALVRNKLFSYEDGHCRCIRKRLPLRALGLRAWTGQR
KDQFGRNSEIPIVQVDFPFGDGGVSLYKNPVPANVEGNDVNFRTMDVPNTL
HAAGVSIIGSEPCPTRAVLPGQHEREGRWWEDAKAKCEGLHKGNIKENTNGNATNV
GTATVADIFNSENVNLSRQGIENLMKLENKRAWIVVLYAPWCPFCQAMEASFDELA
DKLGGSGVKVAKFRADGQDQFAKLEQLGSPFTILVFPKNSRPPIKYPSEKRDVDSL
TSFLNLVR"

CDS

ORIGIN

Query Match 40.3%; Score 657.2; DB 15; Length 1702;
Best Local Similarity 69.5%; Pred. No. 3.2e-162;
Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;
QY 31 AATGGCCCTCGCTTTCAGTCTTCAATTTCCGACCAACTTCCACTTCCCATCATCGGA 90
Db 89 AATCAACGTTTCTTCATCTCTTCTTCTGCGATCTCAAGCTAGTCTTCTTCTTCTCAGA 148
QY 91 ACCCAAACTTCGCGAAATTTGGTCAATTTAGGATTTTCGGAGAGGCCCATTTGGAGGCCCT 150

Db 149 TCTCAAGTAAACAAAAATCGGATCATTTGAGGTTATTTGAATCGTACCAATGTTCTCTGCGGC 208
 Qy 151 TAAATTTCAATTTATCTCAAGACCGGAGCTTGGTAAAGCCCGTTAAACGCCGCAACCTCCACG 210
 Db 209 TTCTCTGAGTTTGTTCGGGAAGAGATCCTCGTGAAGCTCTTAATGTGCATCAATTAC 268
 Qy 211 CAAGGATTCATTTGTTCTCTCGCAGCAACCAATCGTTGCTTCTGTTCTGAGACGAA 270
 Db 269 AAAGGAATCCATTTGTTCTCTGAGGTTACAGAAAGCTAG-----ATGTGTGGA 319
 Qy 271 AGAGGAGATTTTCAACAGATAGCCAGTGTCTCGCAATGCTTCACTCTTGAATCAT 330
 Db 320 AGTTGAAGACTTTCAGGAACCTAGCAAGAGATTAGAGAAATGCTTCTCTCTTGAATCAT 379
 Qy 331 GGATAGAGCCCTCGACAAATTCGCAACGACATAGCTATTGCTTCACTGCTGCTGAAGA 390
 Db 380 GGATAAGACTTTCAGGAAGTTTGGGAATGACATGCTATTGCTTTAGTGAGCTGAAGA 439
 Qy 391 TGTGCTTTGATGATGCGAAATTTGACGGGTGCGACCTTTTAGGCTTTTCACTTTTGA 450
 Db 440 TGTGCTCTGATGATGCTCATTTAACTGGAAGACCTTATAGGCTATTTCAGTTTGA 499
 Qy 451 CACTGAGAGCTGAACCCAGAACTTATCACTTTTGTGTCGGTTGAGAGCAATTTAGG 510
 Db 500 TACAGGGAGATTGAATCCAGAAACATATAGACTCTTCGTACCGTGGAGAAACATTAACG 559
 Qy 511 AATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAGGTTCAAGCATTTGGTGAGAGTAA 570
 Db 560 TATTGCGATTGAGTATGTTCTCTGATGCTGTTGAGGTTCAAGCTTTAGTTAGGAACAA 619
 Qy 571 GGGGTTATCTCTTTTACAGGATGGCCACCAAGAGTGTGTCAGGTTGAGAAAGGTGAG 630
 Db 620 GGGTTGTTCTTCTTCTATGAAGATGTCACACGAGGTTGTCGCTATTAGAAAGGTGAG 679
 Qy 631 GCCTTTAAGGAGGCCCTTAAAGGTTCTCAGAGCATGATTAAGTCTCAGAGGAAGACCA 690
 Db 680 ACCTTTGAGGCGTGTGAAGGGTTTACGTGCTTGGATCACTGCTCAGAGGAAGATCA 739
 Qy 691 GTCACTTGTTACTAGTGTCAAAATACCGGTTGTTTCAAGTGTGATCCGGCTTTTGAAGGAAT 750
 Db 740 ATCAACCGGACAGATCGGAGATTCCTGTTGTTCAAGTGTGATCCAGTGTTTGAGGATT 799
 Qy 751 GGAATGTTGGAATTTGGAAGCTTGTGGAAGTGAACCCCTGTTGCAAAATGTAAGGCCCATGA 810
 Db 800 AGATGGTGGAGTTGGTAGTTGGTGAAGTGAATCCGGTTGCTAAATGTTGAAGGGAATGA 859
 Qy 811 CATATGGAATCTCTTAGGACCATGAATGTCCTGTGGAATTCCTTGCATGCAAAAGGATA 870
 Db 860 TGTTTGAATTTCTTGAGGACGATGGATGTTCTGTGTCAACACCGCTTCACGCTGCGGGTTA 919
 Qy 871 TGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGAAGGGA 930
 Db 920 TGTTTCAATCGGATGTGAGCATGACAGAGCGGTTTTCGGGTGTCAGCATGAGAGA 979
 Qy 931 AGGAGGTGTGTGGGAGGATGCAAGAGCTTAAGGAAATGTTGTTTCAACAAAGGAATGT 990
 Db 980 AGGAGATGTGTGGGGAAGATGCTAAGGCTAAAGAGTGTGGCTTTCACAAAGGGAATAT 1039
 Qy 991 AAAGCAGCAAGAGAGAGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGG 1050
 Db 1040 CAAGGAGAAATCCAAATGGAACCGAACCCGCTAATG-----TCAA 1078
 Qy 1051 TGATGCTPACCACCTGTGCTGCATTTTCAACAGCCCGAATGTAGTTAACTTGAAGCAGGAC 1110
 Db 1079 CGGACAGCCACAGTCGCGGATATCTTCAATAGCGAAGATTTGTGTAACCTTGAAGCAGGCA 1138
 Qy 1111 TGGAAATTGAAATTTGCAAAATTTGAGGAGCCGAAAGAACCAATGCTGTTGTTGCTTTA 1170
 Db 1139 AGGGATTGAAATTTGATGAATTAGAGAACCGGAAGAGGCTTGGATCGTTGTTGCTTTA 1198
 Qy 1171 TGACCATGTGTCCTTACTGCGAGGCTATGAGGAATCTTATGTTGATGTTAGTACGACAA 1230

Db 1199 TGCACCTGTGTGCCGCTTTTGTCTCAGCAATGGAAGCTTCATTCGATGAATTTGGCGGATA 1258
 Qy 1231 GTTAGCAGGCTCAACAGGATGAAGTTTGGAAAATTTAGAGCAGATGAGAAACAGAAAGA 1290
 Db 1259 GCTGGGAGG--AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGGTGACCAAGAGA 1315
 Qy 1291 ATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATATTTTTTCCCAAGCA 1350
 Db 1316 CTTTGCCAAAAGAGGTTGCAGCTTGGGAAGCTTTCACAACTACTTGTGTTTCCCAAGAA 1375
 Qy 1351 TTGCTCTCGGCCCAACATAAAGTATCCCTCAGAAAAGAGAGATTTGATTCCTTGATGGC 1410
 Db 1376 CTCTTCAAG--ACCAATCAAGTATCCATCAGAGAAGAGATGTTGATTCCTTTGACATC 1432
 Qy 1411 ATTTGTAAT 1420
 Db 1433 TTCTTGAAT 1442

RESULT 14 AF428445

LOCUS AF428445 1709 bp mRNA linear PLN 18-OCT-2001
 DEFINITION Arabidopsis thaliana AT4g21990/Fin20_90 mRNA, complete cds.
 ACCESSION AF428445
 VERSION AF428445.1 GI:16226613
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE AUTHORS

1 (bases 1 to 1709)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
 Arabidopsis cDNA clones
 Unpublished
 2 (bases 1 to 1709)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE JOURNAL REFERENCE AUTHORS

Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

TITLE JOURNAL

COMMENT

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shim,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)

contributed equally to this work as PIs.
 Location/Qualifiers
 1. .1709
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="RAPL09-71-K05 (R14914)"
 /ecotype="Columbia"
 1. .84
 85. .1461
 /note="PRH26 protein"
 /codon_start=1
 /product="AT4g21990/Fin20_90"
 /protein_id="AAL16214.1"
 /db_xref="GI:16226614"
 /translation="MALAINVSSSSSAISSSPSSDLKVTIKIGSLRLRLNRNVSAA
 SLISGRKSSVKALNVQSIKESIVASEVTEKLDVVEVDPEELAKRLLENASPLEIMH
 KALEKFNDAIAFSGAEDVALIEYAHLTGRPYRVPFLDTRGLNPETVRLPDTVEKH
 GRIEYPPDAVEQVALVRNKGFLSFYEDGHQBCCRIRKVRPLRLALKGLRMITGQR
 KDQSPGTRSEIPVQVDPVDFEGDGVGSLVKNPVAANVEGNDVNFRLTMDVFPNTL
 HAAGYVIGPEPTRAVLPGQHEREGRWMEKAKCEGLHKGNIKENTNGNATFANVL
 GTASVADIFNSENVLSROGIENLKLNRKEAWIVLVYAPWPCFQAMEASFDELA
 DKLGGSGVKAKFRADQDQDFAKKELQLGSFPTILVFPKNSSRPVKYPSKRDVDSL
 TSEFLNLR"
 1462. .1709
 3'UTR
 misc_difference 1675^1676
 /note="compared to genome"
 /replace="t"
 ORIGIN
 Query Match 40.3%; Score 657.2; DB 15; Length 1709;
 Best Local Similarity 69.5%; Pred. No. 3.2e-162;
 Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;
 31 AATGCCCTCGCTTCACTTCTCAATTTCCGACCACTTCCACCTTCCCATCATCGGA 90
 96 AATCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 155
 91 ACCCAAACTCCGCAAAATGGGTCAATTTAGGATTTCCGAGAGGCCATTTGAGAGCGCGT 150
 156 TCTCAAGTACAAAATCGGATCATTTAGGTTATTTGAATGTACCAATGCTCTCGCGC 215
 151 TAAATTCATTTATCTCAAGAGGAGCTTGGTAAAGCCGTTAAAGCCGGAACCTCCAG 210
 216 TTCTCTGAGTTGTCCGGGAAGAGATCTCTCGGTGAAAGCTCTTAATGTGCAATCAATTAC 275
 211 CAAGGATTCATTTCTCTCGAGCAACCAACCATCGTTCTTCTGCTTCTGAGACGAA 270
 276 AAAGGAATCCATTTGTTCTTCTGAGTTACAGAGAGCTAG-----ATGGTGGGA 326
 271 AGAGGAAGATTTTGAACAGATAGCCAGTGTCTCGACAATGCTTCCACCTCTTGAATCAT 330
 327 AGTTGAAGACTTTGAGGAACCTAGCAAGAGATTTAGAGAATGCTTCTCTCTTGAATCAT 386
 331 GGATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGCTTCACTGCTGCTGAAGA 390
 387 GGATAAAGCTCTTGAGAAAGTTGGGAATGACATTTGCTATTGCTTTAGTGGAGCTGAAGA 446
 391 TGTGCTTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 450
 447 TGTGCTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 506
 451 CACTCGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTGAGAGCAATTTATGG 510
 507 TACAGGAGATTTGAATCCAGAAACATATAGACTCTTCGATACCGTGGAGAAACATTTACG 566
 511 AATTCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 570
 567 TATTCGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 626
 571 GGGGTTATTTCTTTCTAGAGGATGGGCAACAGAGTGTTCAGGGTGAGAAAGGTGAG 630

Db 627 GGCTTTGTTCTCTTTCTATGAAGATGGTCAACAGAGTGTTCGCGTATTAGAAAGGTGAG 686
 Qy 631 GCCTTTAAGAGAGCGCCCTTAAGGGTCTCAGAGCATGGAATACTGGTCAGAGGAAGACCA 690
 Db 687 ACCTTTGAGCGCGTGGTGAAGGGTTTACGTGCTTGGATCAGTGTGAGAGGAAGATCA 746
 Qy 691 GTCACTGTGTACTAGGTCTCAAAATACCGGTGTGTTCAGTGTGATCCGCTTTTGAAGGAAT 750
 Db 747 ATACCCGGGACAGATCGGAGATTCCTGTTGTTCAAGTTGATCCAGTGTGTTGAAGGATT 806
 Qy 751 GGATGTGGAATTTGGAAGCTTGTGGAAGTGGAAACCCCTGTTGCAAAATGTGAAGGCCATGA 810
 Db 807 AGATGGTGGAGTTGGTAGTTGGTGAAGTGGAAATCCGGTTGCTAAATGTTGAAGGGAATGA 866
 Qy 811 CATATGGAATCTCTTAGGACCATGAATGTCCTGTGCAATTCCTTGCATGCAAAAGGATA 870
 Db 867 TGTTTGGAATTTCTTAGGACCATGAATGTTCTGTCTCAACACGCTTCACGCTCGGGGTTA 926
 Qy 871 TGTTCATTTGGGTGTGAGCCCTGCACTAGCGCTTGTCTTACCTGGGCAACATGAAAGGGA 930
 Db 927 TGTTCATCGGATGTGAGCATGCAAGAGCGGTTTTCGCGGTGAGCATGAGAGAGA 986
 Qy 931 AGGAGGTGTGTGGGAGGATGCAAGCTTAAGGAATGTGCTTCAAAAGGAATGT 990
 Db 987 AGGAGATGTGTGGGAGAGATGCTAAGGCTTAAAGAGTGTGGCCTTCAAAAGGGAATAT 1046
 Qy 991 AAAGCAGAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAAATGG 1050
 Db 1047 CAAGGAGAAATCAACATGGAACCGACCGCTAAATG-----TCAA 1085
 Qy 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTTGAGCAGGAC 1110
 Db 1086 CGGAGAGCCCTCAGTCGGGATATCTTCAATAGCGAGATGTTGTGAATTTGAGCAGGCA 1145
 Qy 1111 TGGAAATTTGCAAAATTTGCAAAATTTGAGGAGCCGAAAGGAACATGCTTGTGCTTTA 1170
 Db 1146 AGGATTTGAGAAATTTGATGAATTTAGAGAACCGGAAAGAGGCTTGGATCGTTGCTTTA 1205
 Qy 1171 TGCACCATGTGCCCCCTACTGCCAGGCTATGAGGAGAACTTTATGTTGACTTTAGCAGACAA 1230
 Db 1206 TGCACCGTGTGCCCCGTTTGTGAGCAATGGAAGCTTCATTCGATGAATTTGGCGGATTA 1265
 Qy 1231 GTTAGCAGGCTCAACAGGAGTGAAGGTTTGAAGAAATTTAGAGCAGATGAGAAACAGAAAGA 1290
 Db 1266 GCTCGGAGG---AAGTGTGTGAAGTTGCTTAAGTTTAGAGCTGATGTCAGCAGAGGA 1322
 Qy 1291 ATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTATTTTCCCAAAGCA 1350
 Db 1323 CTTTGCCAAAAGAGGTTGCAGCTTGGAAAGCTTTCCAAACAATACCTGTGTGTTCCCAAAGAA 1382
 Qy 1351 TTGCTCTCGGCAACAAATAAGATCCCTCAGAAAGAGAGATGTTGATTTCCCTTCATGCGC 1410
 Db 1383 CTCTTCAAG---ACCAATCAAGTATCCATCAGAGAAGAGAGATGTTGATTTCTTGACATC 1439
 Qy 1411 ATTTGTAAT 1420
 Db 1440 TTTCTTGAAT 1449

RESULT 15

AY054175

LOCUS

Arabidopsis thaliana

AT4g21990/Fin20_90

mRNA

linear

PLN 30-SBP-2001

DEFINITION

Arabidopsis thaliana

AY054175

ACCESSION

AY054175.1

VERSION

FLI CDNA.

KEYWORDS

Arabidopsis thaliana

SOURCE

Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1712)

AUTHORS

Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banth, J.,

Db	1147	AGGATTGAGAA	TTGATGAAGTTAGAGAACCGGAAGAGGCTTGGATCGTTGTGCTTTA	1206
Qy	1171	TGCACCATGGT	TGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGCAA	1230
Db	1207	TGCACCGTGGT	TGCCCGTTTGTTCAGGCAATGGAAGCTTCATTCGATGAATTTGGCGGATAA	1266
Qy	1231	GTTAGCAGGTC	AACAGGGATGAAGTTGGAAATTTTAGACGAGATCGAGAACAGAAAGA	1290
Db	1267	GCTGGGAGG---	AAGTGGTGTGAAGTTGCTAAGTTTAGAGCTGATGGTGACCCAGAAAGGA	1323
Qy	1291	ATTGCAAGAGT	GAACTGCAATTGGGAAGCTTCCCTACGATATTATTTTCCCAAAGCA	1350
Db	1324	CTTTGCCAAA	AAGAGTTGCAGCTTGGAGCTTTCCAACTACTTGTGTTCCCAAAGAA	1383
Qy	1351	TTGCTCTCGGC	CAACAATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTTGATGGC	1410
Db	1384	CTCTTCAAG---	ACCAATCAAGTATCCATCAGAGAAGAGAGATGTTGATTCTTTGACATC	1440
Qy	1411	ATTGTAAAT	1420	
Db	1441	TTTCTGAAT	1450	

Search completed: February 20, 2006, 23:16:02
 Job time : 8433 secs


```

Db      1381 CAACCAATTAAGTACCCTTCGGAAAAAGAGAGATGTTGATTCAATGACGGCAATTCGTGAAT 1440
Qy      1421 GCCTTAAGATGAGAT 1436
      |||||
Db      1441 GCCTTAAGATGAGAT 1456

RESULT 3
AAZ50065
ID      AAZ50065 standard; cDNA; 1795 BP.
XX
AC      AAZ50065;
XX
DT      04-MAY-2000 (first entry)
XX
DE      Soybean adenosine-5'-phosphosulfate reductase cDNA clone se4.11g09.
XX
KW      Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;
KW      sulphate assimilation protein; SAP; clone se4.11g09; ss.
XX
OS      Glycine max.
XX
FH      Key
FT      CDS      Location/Qualifiers
FT          50..1468
FT          /*tag= a
FT          /product= "Soybean APS reductase"
XX
XX      WO200004161-A1.
XX
XX      27-JAN-2000.
XX
XX      13-JUL-1999; 99WO-US015808.
XX
XX      14-JUL-1998; 98US-0092833P.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Falco SC, Allen SM;
XX
XX      WPI; 2000-182427/16.
XX      P-PSDB; AAY44744.
XX
XX      Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX      and seeds useful as probes for isolating cDNAs and genes encoding
XX      homologous proteins, and in producing transgenic plants.
XX
XX      Claim 3; Page 32-33; 47pp; English.
XX
XX      The present cDNA sequence encodes soybean adenosine-5'-phosphosulfate
XX      (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is
XX      designated se4.11g09, isolated from soybean embryo, 19 days after
XX      flowering. APS reductase polypeptides are useful for producing antibodies
XX      for screening cDNA expression libraries to isolate full length APS
XX      reductase cDNA clones. Chimeric genes comprising the APS reductase gene
XX      when transformed in a host cell can alter the level of SAP expression.
XX      This can be used to increase sulphur containing compounds in the cell,
XX      including the nutritionally critical amino acids cysteine and methionine
XX
XX      Sequence 1795 BP; 504 A; 351 C; 436 G; 504 T; 0 U; 0 Other;
XX
XX      Query Match      55.5%; Score 904; DB 3; Length 1795;
XX      Best Local Similarity 81.3%; Pred. No. 9.4e-250;
XX      Matches 1103; Conservative 0; Mismatches 235; Indels 18; Gaps 4;
XX
Qy      81 CATCATCGGAACCCAAACTTCGGCAAAATGGGTCAATTAGGATTTTCGGAGAGGCCCATTTG 140
Db      135 CTTTCATCGGACGCTAAGCTCGCAAAATGGTTCCTTCGGTTCCTCGGAGAGGCGCTCAAG 194
Qy      141 GAGGCGCGTTAAATTTCAATTTATCTCAAGACGGAGCTTGGTAAAGCCGCTTAACGCCG 200
Db      195 TTTCTGCTGGTGTGTTAAATTTAACTCAAGACGCTCTCTCGGTGAGGCCACTCAATGCCG 254
Qy      201 AACCTCCAGCAAGGATTCATTTGTTCTCTCGCAACCAACCATTCGTTGCTTCTGCTT 260

```

```

Db      255 AACCGCAACGGAATGATTCCTGTTGTTCTCTTGCGAG---CAACTATGTTGCTCTCTGAGG 311
Qy      261 CTGAGACGAAAGAGGAATTTTGAACAGATAGCCAGTATCTCGAAGTCTTCACTTC 320
      |||||
Db      312 TTGAGAAAGGAGAAAGAAAGATTTTGAAGCAATTTAGCGAAAGACCTTTGAAATTCATCTCCTC 371
      |||||
Qy      321 TTGAAATCATGGATAGAGCCCTCGACAAATTTGGCAACAGACATAGCTATTATTCCTTCAGTG 380
      |||||
Db      372 TTGAGATTTGATAGGCCCTCGAGAAATTTGGGAACGACATCGCTATTGCCCTTTAGTG 431
      |||||
Qy      381 GTGCTGAAGATGTTGCTTTGATTGAGTATGCGAAATTTGACGGGTGCGACCTTTTAGGGTTT 440
      |||||
Db      432 GTGCTGAAGATGTTGCTTTGATTGAGTATGCGAATTTTGAACGGGTGCGACCTTCACAGAGTGT 491
      |||||
Qy      441 TCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCACTTTTGTGATCGGTTGAGA 500
      |||||
Db      492 TTAGTCTTGACACTGGGAGACTGAACCCAGAACTTCAAAATTTTGGACCTGTTGAGA 551
      |||||
Qy      501 AGCATTATGGAATTCGCATTTGAGTACATGTTCCCTGATGCTGTTGAGGTTGAGGTTGAGGTTAG 560
      |||||
Db      552 AGCATTATGGAATTCATATTGAGTACATGTTCCCTGATGCGGTTGAGGTTGAGGTTAG 611
      |||||
Qy      561 TGAGGAGTAAAGGGGTTATTTCTTTTCTACGAGGATGGGCACCAAGAGTGTGTCAGGGTGA 620
      |||||
Db      612 TAAGAACTAAGGGGCTCTTCTCTATTTTACGAGGATGGGCATCAAGAGTGTGTAGAGTAA 671
      |||||
Qy      621 GAAAGGTGAGGCTTTTAAAGGAGGCGCTTAAAGGCTCTCAGAGCATGATTAATCTGTCAGA 680
      |||||
Db      672 GAAAGGTGAGGCGCTTTGAGGAGAGCGCTTAAAGGCTCTCAAGAGCATGGAATTAATCTGACAGA 731
      |||||
Qy      681 GGAAGACCGAGTCACTGAGTCTAGGTTGAGTCTGAAATACCGGTTGTTTTCAGGTTGATCCGGCTT 740
      |||||
Db      732 GAAAGACCGAGTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 791
      |||||
Qy      741 TTGAGGGAATGGATGGTGGAAATTTGGAAGCTTTGGTGAAGTGGAAACCTGTTGCCAATGTGA 800
      |||||
Db      792 TTGAGGAGCTGGATGGTGGAAATTTGGCAGCGCTGGTGAAGTGGAAACCTGTTGCCAATGTGA 851
      |||||
Qy      801 AGGCGCATGATGATATGGAACCTTCTTAGGACCATGAAATGTCCTGTGAAATTCCTTGCATG 860
      |||||
Db      852 ATGGTCTAGACATATGGAACCTTCTTAGGACCATGAAATGTTCTCTGTAATTCATTTGCAAT 911
      |||||
Qy      861 CAAAGGATATGTTTCCATTTGGGTGAGCGCTGCACCTAGGCGCTGTTTACCTGGGCAAC 920
      |||||
Db      912 CCAAGATATGTTTCGATTTGGCTGTGAGCCATGCAACAGGCGCGTTTACCCGACAC 971
      |||||
Qy      921 ATGAAAGGGAAGGAGGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGGTCTTTCA 980
      |||||
Db      972 ATGAAAGGAAGGAAGGTGGTGGTGGGAGGATGCCAAAGCCAAAGAGTGTGGTCTTTCA 1031
      |||||
Qy      981 AAGGAATGTTAAAGCAGCAGAAAGAGGAGATGTTTATGGAATGGGCTATCCCAATCCC 1040
      |||||
Db      1032 AAGGTAATTTGAAACAGGAAGATGCTGCCAGCTTAATGSAATGGGACCTCCCAAGGAA 1091
      |||||
Qy      1041 ATGCAATGGTGTGATGTACCACTGTGCTGACATTTTCAACAGCGCCGAATGTAGTTAACT 1100
      |||||
Db      1092 A-----TGGCTCTGCCACTGTTGCTGACATTTTCTCCAGAAATGTGGTCACT 1142
      |||||
Qy      1101 TGAGCAGGACTCGAAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGAACCATGGCTTG 1160
      |||||
Db      1143 TGAGCAGGTCCGGGATTTGAGAAATTTGGCAAAATTTAGAGAACCGAAAGAACACATGGCTTG 1202
      |||||
Qy      1161 TTGTGCTTTATGACCATGTTGCCCTTACTGCGAGGCTATGAGGAATCTTATGTTGACT 1220
      |||||
Db      1203 TTGTGCTTATGACCATGTTGCCCTTCTGTGAGGCTATGAGGAGTGTGATGTTGATC 1262
      |||||
Qy      1221 TAGCAGACAAAGTTAGCAGGCTCAACAGGATGAAAGTTGGAATTTTAGAGCAGATGAG 1280
      |||||
Db      1263 TGGCAGAGAGTTTAGCAAGGT---CAGAGTGAAGTTGCAAAATTCAGAGCCGATGAG 1319
      |||||
Qy      1281 AACGAAAGAAATTTGCAAGAGTGAATGCAATGGGAAGCTTCCTCAGATATTTATTTT 1340
      |||||

```

Db	1320	ACGAGAAGGAATATGCAAGAAGTGAACGACGTTGGGAGCTTCCCAACATACTTCTCT	1379	PR	10-JUN-1999;	99US-0138540P.
Qy	1341	TCCCAAGCAATCGTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGAGATCTTGATT	1400	PR	10-JUN-1999;	99US-0138847P.
Db	1380	TCCCAAGCACTCTTCT--CAACCAATTAAGTACCCTTCAGAAAAGAGAGATCTTGATT	1436	PR	14-JUN-1999;	99US-0139119P.
Qy	1401	CTTTGATGGCAATTTGTAATGCGCTTAAGATGAGGAT	1436	PR	16-JUN-1999;	99US-0139452P.
Db	1437	CATTGACGGCATTCTGTAATGCTTACGGTGATGGT	1472	PR	17-JUN-1999;	99US-0139453P.
RESULT 4				PR	18-JUN-1999;	99US-0139454P.
AAC43415				PR	18-JUN-1999;	99US-0139455P.
ID AAC43415	standard; DNA; 1377 BP.			PR	18-JUN-1999;	99US-0139456P.
XX				PR	18-JUN-1999;	99US-0139457P.
AC AAC43415;				PR	18-JUN-1999;	99US-0139458P.
XX				PR	18-JUN-1999;	99US-0139459P.
DT 17-OCT-2000	(first entry)			PR	18-JUN-1999;	99US-0139460P.
XX				PR	18-JUN-1999;	99US-0139461P.
DE Arabidopsis thaliana	DNA fragment SEQ ID NO: 39173.			PR	18-JUN-1999;	99US-0139462P.
XX				PR	18-JUN-1999;	99US-0139463P.
KW	Hybridisation assay; genetic mapping; gene expression control;			PR	18-JUN-1999;	99US-0139750P.
KW	protein identification; signal transduction pathway; metabolic pathway;			PR	18-JUN-1999;	99US-0139763P.
KW	promoter; termination sequence; ss.			PR	21-JUN-1999;	99US-0139817P.
XX				PR	22-JUN-1999;	99US-0139899P.
OS Arabidopsis thaliana.				PR	23-JUN-1999;	99US-0140353P.
XX				PR	23-JUN-1999;	99US-0140354P.
PN EP1033405-A2.				PR	24-JUN-1999;	99US-0140895P.
PD 06-SEP-2000.				PR	28-JUN-1999;	99US-0140823P.
XX				PR	29-JUN-1999;	99US-0140991P.
XX				PR	30-JUN-1999;	99US-0141287P.
XX				PR	01-JUL-1999;	99US-0141842P.
XX				PR	01-JUL-1999;	99US-0142154P.
XX				PR	02-JUL-1999;	99US-0142055P.
XX				PR	06-JUL-1999;	99US-0142390P.
XX				PR	08-JUL-1999;	99US-0142803P.
XX				PR	09-JUL-1999;	99US-0142977P.
XX				PR	12-JUL-1999;	99US-0143542P.
XX				PR	13-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	99US-0121825P.		PR	14-JUL-1999;	99US-0143624P.
PR	05-MAR-1999;	99US-0123180P.		PR	15-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	99US-0123548P.		PR	16-JUL-1999;	99US-0144085P.
PR	23-MAR-1999;	99US-0125788P.		PR	16-JUL-1999;	99US-0144325P.
PR	25-MAR-1999;	99US-0126264P.		PR	19-JUL-1999;	99US-0144331P.
PR	29-MAR-1999;	99US-0126785P.		PR	19-JUL-1999;	99US-0144332P.
PR	01-APR-1999;	99US-0127462P.		PR	19-JUL-1999;	99US-0144333P.
PR	06-APR-1999;	99US-0128234P.		PR	19-JUL-1999;	99US-0144335P.
PR	08-APR-1999;	99US-0128714P.		PR	20-JUL-1999;	99US-0144352P.
PR	16-APR-1999;	99US-0129845P.		PR	20-JUL-1999;	99US-0144632P.
PR	19-APR-1999;	99US-0130077P.		PR	20-JUL-1999;	99US-0144884P.
PR	21-APR-1999;	99US-0130449P.		PR	21-JUL-1999;	99US-0144814P.
PR	23-APR-1999;	99US-0130510P.		PR	21-JUL-1999;	99US-0145086P.
PR	23-APR-1999;	99US-0130891P.		PR	21-JUL-1999;	99US-0145088P.
PR	28-APR-1999;	99US-0131449P.		PR	22-JUL-1999;	99US-0145085P.
PR	30-APR-1999;	99US-0132048P.		PR	22-JUL-1999;	99US-0145087P.
PR	30-APR-1999;	99US-0132407P.		PR	22-JUL-1999;	99US-0145089P.
PR	04-MAY-1999;	99US-0132484P.		PR	22-JUL-1999;	99US-0145192P.
PR	05-MAY-1999;	99US-0132485P.		PR	23-JUL-1999;	99US-0145145P.
PR	06-MAY-1999;	99US-0132486P.		PR	23-JUL-1999;	99US-0145218P.
PR	06-MAY-1999;	99US-0132487P.		PR	23-JUL-1999;	99US-0145224P.
PR	07-MAY-1999;	99US-0132863P.		PR	25-JUL-1999;	99US-0145276P.
PR	11-MAY-1999;	99US-0134256P.		PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134218P.		PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134219P.		PR	27-JUL-1999;	99US-0145919P.
PR	14-MAY-1999;	99US-0134221P.		PR	28-JUL-1999;	99US-0145951P.
PR	18-MAY-1999;	99US-0134370P.		PR	02-AUG-1999;	99US-0146386P.
PR	18-MAY-1999;	99US-0134768P.		PR	02-AUG-1999;	99US-0146388P.
PR	19-MAY-1999;	99US-0134941P.		PR	02-AUG-1999;	99US-0146389P.
PR	20-MAY-1999;	99US-0135124P.		PR	03-AUG-1999;	99US-0147038P.
PR	21-MAY-1999;	99US-0135353P.		PR	04-AUG-1999;	99US-0147204P.
PR	24-MAY-1999;	99US-0135629P.		PR	04-AUG-1999;	99US-0147302P.
PR	25-MAY-1999;	99US-0136021P.		PR	05-AUG-1999;	99US-0147192P.
PR	27-MAY-1999;	99US-0136392P.		PR	05-AUG-1999;	99US-0147260P.
PR	28-MAY-1999;	99US-0136782P.		PR	06-AUG-1999;	99US-0147303P.
PR	01-JUN-1999;	99US-0137222P.		PR	09-AUG-1999;	99US-0147416P.
PR	03-JUN-1999;	99US-0137528P.		PR		99US-0147493P.
PR	04-JUN-1999;	99US-0137502P.		PR		
PR	07-JUN-1999;	99US-0137724P.		PR		
PR	08-JUN-1999;	99US-0138094P.		PR		


```
Db 1062 AGGATTGAGAAATTTGATGAAGTTAGAGAACCGGAAAGAGCGCTGGATCGTTGTGCTTTA 1121
Qy 1171 TGCACCATGGTGCCTACTCCAGGCTATCGAGGAATCTTATGTTGACTTTAGCAGACAA 1230
Db 1122 TGCACCGTGGTGCCTGTTTGTGAGGCAATGGAAGCTTCATTGATGAATTTGGCGGATAA 1181
Qy 1231 GTTAGCAGGTCAACAGGATGAAGTTGGAAATTTAGCAGATGGAGAACAGAAAGA 1290
Db 1182 GCTGGGAGG---AAGTGGTGTGAAGGTTGCTTAAGTTTAGAGCTGATGGTGACCAAGAAG 1238
Qy 1291 ATTTGCCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTTATTTTCCCAAGCA 1350
Db 1239 CTTTGCCAAAAGAGTTGAGCTTGGAGCTTTCCAACTACTTGTGTTCCCAAGAA 1298
Qy 1351 TTGCTCTGGCCAAACAATAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTTGATGGC 1410
Db 1299 CTCCTTCAAG---ACCAATCAAGTATCCATCAGAGAAGAGAGATGTTGATTCCTTTGACATC 1355
Qy 1411 ATTTGTAAT 1420
Db 1356 TTTCTTGAAT 1365

RESULT 5
AAC38867
ID AAC38867 standard; DNA; 1702 BP.
XX
AC AAC38867;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 22518.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135153P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
```


PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147433P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 40.3%; Score 657.2; DB 3; Length 1702;		
Best Local Similarity 69.5%; Pred. No. 1.6e-178;		
Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;		
Qy	31	AATGGCCCTCGCTTCACTTCTTCAATTCGGCACCAACTTCCACCTTCCCATCATCGGA 90
Db	89	AATCAACGTTTCTTCATCTTCTTCTCTGCGATCTCAAGCTCTAGCTTCCCTTCTTCAGA 148
Qy	91	ACCCAACTTCGCGAAATTTGGTCAATTAGGATTCGGAGAGGCCCATTCGAGGGCGGT 150
Db	149	TCTCAAGTAAACAAAATCGGATCAATGAGGTTATTAATGCTGATCAATGCTCTCGGCG 208
Qy	151	TAATTTCAATTTATCTCAAAGACGGAGCTTGGTAAAGCCGTTAAACGCCGAACCTCCACG 210
Db	209	TTCTCTGAGTTTGTCCGGAGAGATCTCTCGTGAAGCTCTTAATGTGCATCAATTAC 268
Qy	211	CAAGGATTCATTTCTCTCTCGCAGCAACCACTCGTTGCTTCTGCTCTGAGACGAA 270
Db	269	AAAGGAATCCATTGTTGCTTCTGAGGTTACAGAGAAGCTAG-----ATGTGGTGA 319
Qy	271	AGAGGAAGATTTTGAACAGATAGCCAGTATCTCGCAANTGCTTCACTCTTGAATCAT 330
Db	320	AGTTGAAGACTTTGAGGAACCTAGCAAGAGATTAGAGAAATGCTTCTCTCTTGAGATCAT 379
Qy	331	GGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCCCTTCAGTGTGCTGAAGA 390
Db	380	GGATAAAGCTCTTGAGAAGTTTGGGAATGACATTTGCTATTGCTTTAGTGAGCTGAAGA 439
Qy	391	TGTTGCTTTGATTCAGTATGCGAAATTCACGGGTCGACCCCTTTAGGGTTTTCAGTTTGA 450
Db	440	TGTTGCTCTGATTCAGTATGCTCATTTAACTGGAAGACCTTATAGGGTATTTCAGTTTGA 499
Qy	451	CACCTGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGCGGTTGAGAGCATTTGG 510
Db	500	TACAGGGAGATTGAATCCAGAAACATATAGACTCTTCGTATCCGCTGGAGAAACATTACGG 559
Qy	511	AATTGCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGCATTGGTGAGAGTAA 570
Db	560	TATTCGGATTGAGTATATGTTTCTGATGCTGTTGAGGTTCAAGCTTTAGTTAGGAACNA 619
Qy	571	GGGGTTATTTCTTTCTACGAGGATGGGCACCAAGAGTGTTCGAGGGTGAGAAAGGTGAG 630
Db	620	GGGTTTGTCTTCTTATGAAGATGGTCAACAGGAGTGTTCGCCGTATTAGAAAGGTGAG 679
Qy	631	GCCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGAATACTGTCAGAGGAAGAGCA 690
Db	680	ACCTTTGAGGGTGGCTTGAAGGGTTTACGTGCTTGGATCACTGGTCAGAGAAAGATCA 739
Qy	691	GTCACCTGCTACTAGTCTGAAATACCGGTTGTTTCAGGTTCATCCGGCTTTTGAGGGAAT 750
Db	740	ATCACCCGGGAACAGATCGGAGATTCCTGTTGTTTCAGGTTGATTCAGTGTTCGAGGATT 799
Qy	751	GGATGGTGGAAATTTGGAAGCTTTGGTGAAGTGGAAACCTCTGTCGAAATGTGAAGGGCATGA 810
Db	800	AGATGGTGGAGTTGGTAGTTTGGTGAAGTGGAAATCCGGTGTGCTAATGTTGAAGGGAATGA 859
Qy	811	CATATGGAACCTTCCTTAGGACCATGAATGTGCTGTGAAATTCCTTGCATGCAAAAGGATA 870
Db	860	TGTTTGGAAATTTCTTGAGGACGATGGATGTTTCTGTCAACGCTTCACGCTCGCGGTTTA 919
Qy	871	TGTTTCCATTTGGGTGTGAGCCCTGACCTAGGCGTGTGTTTACCTGGGCAACATGAAGGGA 930
Db	920	TGTTTCAATCGAATGTGAGCCCATGCAACGAGCGGTTTTCGCGGTCAGCATGAGAGAGA 979
Qy	931	AGGGAGGTGGTGGTGGGAGGATGCCAAAGCTTAAGGAAATGTGTCTTTCACAAAGGAATGT 990

```
Db 980 AGGAGATGGTGGGGAAGATGCTAAGGCTAAAGAGTGTGGCCCTTCACAAAGGAATAT 1039
Qy 991 AAACAGCAGAAAAGAGAGAGATGTTAATGGAAATGGGCTATCCCAATCCCATGGCAATGG 1050
Db 1040 CAAGGAGATACCAATGGAAACGCAACCGCTAATG-----TCAA 1078
Qy 1051 TGATGCTACACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGGACGAGAC 1110
Db 1079 CGGACAGCCACAGTCGCGGATATCTTCAATAGCGAGAATGTTGTGAACCTTGGACGGCA 1138
Qy 1111 TGGAAATTGAGAATTTGGCAAAATTTGGAGGACCGAAGAACCATGGCTTGTGTGCTTTA 1170
Db 1139 AGGAATTGAGAAATTTGATGAATTTAGAGAACCGGAAGAGCGCTTGGATCGTTGTGCTTTA 1198
Qy 1171 TGCAACCATGTGTCCTCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTTAGCAGCAA 1230
Db 1199 TGCACCGTGTGTCCTGTTTGTGAGGCAATGGAAGCTTCATTCGATGAATGGCGGATAA 1258
Qy 1231 GTTAGCAGGTCACAGCGATGAAGTTGGAAATTTAGACGAGATGGAGAACAGAAAGA 1290
Db 1259 GCTGGAGG---AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGTCACCAAGGA 1315
Qy 1291 ATTTGCAAGAGTCAATGCAATTTGGGAAGCTTCCCTACGATATATTTTCCCAAGCA 1350
Db 1316 CTTTGGCCAAAAGAGATTGGAGCTTGGAAAGCTTTCCCAACATACTGTGTGTTCCCAAGAA 1375
Qy 1351 TTCTGTCTGGCCAAACAATAAGTATCCCTCAGAAAAGAGAGATGTCATTCCTTGATGGC 1410
Db 1376 CTCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGATGTTGATCTTTTGACATC 1432
Qy 1411 ATTTGTAAT 1420
Db 1433 TTTCTTGAAT 1442

RESULT 6
AAC47286
ID AAC47286 standard; DNA; 1712 BP.
AC AAC47286;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 53259.
DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
PN
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
```

PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	23-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145951P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148331P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154033P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159333P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	22-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161932P.
PR	29-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 40.3%; Score 656.6; DB 3; Length 1712;		
Best Local Similarity 70.8%; Pred. No. 2.4e-178; Mismatches 359; Indels 42; Gaps 6;		
Matches 974; Conservative 0;		
Qy	83	TCATCGGAACCCAACTTCCGCAAAATTGGGTCAATTAGGATTTCCGAGAGCCCAATCGGA 142
Db	133	TCATTGGAGCCAAAGTTTCGCAAAATTGGTTCGTTGAGGTTATTGGATCGTGTTCATGTT 192
Qy	143	GGGCGCGTTAATTCAATTATCTCAAGACGG---AGCTTGGTAAGACCCCGTTAACGCC 199
Db	193	GCTCCTGTCTCTGAAATCTATCTGGGAAGCGATCATCATCTGTAAACCTTTAAACGCT 252
Qy	200	GAACCTCCAGCAAGGATTCATTGTTCTCTCCAGCAACCAACCATCGTTGCTTCTGCT 259
Db	253	GAACCAAGACAAAGGATTCATGATTCCTCTCGGCAACCAATGCTAGAGAAATGCA 312
Qy	260	TCTGAGACGAAGAG-----GAAGATTTTGAACAGATAGCAGTGATCTCGACAAT 310
Db	313	GAGGAAGTTGAAGTGGTTGAGATTTTGAAGAGCTTGTCTAAGAAGTTAGAGAAT 372
Qy	311	GCTTCACCTCTTGAATTCATGAGTAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
Db	373	GCTTCACCTCTTGAGATTTATGGAACAAAGCTTTGAGAAATACGGGAACGATATGCCATT 432
Qy	371	GCCTTCAGTGTCTGGAAGATGTTGCTTTGATTGAGTATGCGAAATTCACGGGTGACGCC 430
Db	433	GCAITTAGTGTGCAAGAGATGTTGCTCTTAITTAGTAGCTCATTTGACTGGGAGGCCA 492
Qy	431	TTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTTGAT 490
Db	493	TTTAGAGTATTTAGTTTGGATACAGGGAGGTTGAATCCTGAGACGTATCGGTTTTTTCGAT 552
Qy	491	CGGTTGAGAGCAATATGGAATTCGCATTTGAGTACATGTTCCCTGATGCTGTTCAGGTT 550
Db	553	CGGTTGAGAGCACTATGGAATTTAGGATTTAGGATTTAGTATATGTTTCTCGATCTGTTCAGGTT 612
Qy	551	CAGCATTTGTTGAGGAGTAAGGGTTATCTCTTTCTACGAGGATGGGCACCAAGAGTGT 610
Db	613	CAAGTTTGGTTAGGACGAGGATGTTCTCTTTTATGAGGATGGTCAACAGAGTGT 672
Qy	611	TGCAGGGTGAGAAAGGTGAGGCCCTTTAAGGAGGGCCCTTAAAGGGTCTCAGAGCATGGATA 670
Db	673	TGCCGTGTTCCGAAGGTGAGACCTTTGAGGGGTGCTCTCAAGGGTTTAAAGGCTTGGATT 732
Qy	671	ACTGTCAGAGGAAGACAGCTACCTGCTAGGTCTGAAATACCCGGTTGTTTCAGGTT 730
Db	733	ACTGTCAGAGGAAGATCAATCTCCGGGCAAGGTCGAGATTCGCGTTGTTTCAGGTT 792
Qy	731	GATCCGGCTTTTGGGGAATGATGCTGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTT 790
Db	793	GATCCGGTGTGGAAGGTTTGGATGGAGGTTGTTAGTGTGTTGTTGTTGTTGTTGTTGTT 852
Qy	791	GCAATGTGAAGGGCCATGACATATGGAACCTTCTTTAGGACCATGAATGTGCTGTGAAT 850

```
Db      853  GCGAATGTTGAAGGAATGATGTTTGGAACTTCTTGAGGACTATGAGATGTTCCGGTTAAC 912
Qy      851  TCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACCTAGGSCCTGTTTTTA 910
Db      913  ACATTGCATGGGAGGATATATCGATTGGATGTGAGCCTTGCGAAGAACGGTTTTTA 972
Qy      911  CCTGGGCAACATGAAGGAAGGAGGTGGTGGTGGAGGATGCGCAAGCTTAAGGAATGT 970
Db      973  CCGGGTCAGCACGAGAGAGGAGATGGTGGTGGGAAGATGCTAAAGCCCAAGGAATGT 1032
Qy      971  GGTCTTCACAAGCAATGTAAGACGACGAGAGGAGGATGTTAATGGAATGGGCTA 1030
Db      1033  GGACTTCACAAGGGAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1089
Qy      1031  TCCCAATCCCATGCAAAATGGTATGCTTACCACTGTGCTGACATTTTCAACAGGCCCAAT 1090
Db      1090  TCGAAT-----CCGCTGTTCGAGATATCTTTAAGATGAGANT 1128
Qy      1091  GTAGTTAACTTGACGAGACTGGAATTGAGAAATTTGCAAAATTTGGAGGACCGAAAGGAA 1150
Db      1129  CTTGTGACTTTGACGAGCGAGGATTTGAGAAATTTGATGAAGTTGGAGAACCGTAAAGAG 1188
Qy      1151  CCATGGCTTGTGCTTTATGACCACTGTTGCCCTTACTGCCAGGCTATGGAGGAATCT 1210
Db      1189  CTTGGATCGTCGCTTTATGCTCCGTGGTGGCCCCCTTTTGTCAAGCCATGGAAGCATCG 1248
Qy      1211  TATGTTGACTTTAGCAGCAAGTTAGCAGGGTCAACAGGATGAAGTTTGGAAATTTAGA 1270
Db      1249  TATGATGAACCTGGGGATNAATTGGCTGG---AAGTGGATTAAGTTGCCAATTCAGA 1305
Qy      1271  GCAGATGAGAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGAAGCTTCCCTTACG 1330
Db      1306  GCAGATGGTGACCAAGAGGAGTTTGCTTAAGCAGGAATTCAGCTCGGTAGCTTCCCTTACC 1365
Qy      1331  ATATTATTTTCCCAAGCATTCGCTCGGCCCAACATAAGTATCCCTCAGAAACAGA 1390
Db      1366  ATCTGTTTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGGAAGAGA 1422
Qy      1391  GATGTTGATTCCTTGATGGCAATTTGTAATGCCCTTAAGATGAGGATATCAGGAAA 1445
Db      1423  GATGTTGAGTCTTGTGACTTGTCTTGAATCTTGTCGTAAGTAAACCAACAAA 1477

RESULT 7
AAC37691
ID AAC37691 standard; DNA; 1714 BP.
XX AC AAC37691;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18317.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      23-APR-1999; 99US-0130891P.
PR      28-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
```

PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144332P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145132P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0146389P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148117P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151068P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	16-SEP-1999;	99US-0154018P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
Query Match 40.2%; Score 655; DB 3; Length 1714;		
Best Local Similarity 70.8%; Pred. No. 7e-178;		
Matches 973; Conservative 0; Mismatches 360; Indels 42; Gaps 6;		
Qy	83	TCATCGGAACCCAACTTCGCAATTTGGGTCAATTAGGATTCGGAGAGCCCAATTGGA 142
Db	133	TCATTGGAGCCAAAGTTTCGCAATTTGGTTCGTTGAGGTTATTGGATCGTGTTCATGTT 192
Qy	143	GGCGCCGTTAATTTCAATTTATCTCAAGACGG---AGCTTGGTAAAGCCCGTTAAACGCC 199
Db	193	GCTCCTGTGCTCTGAACTATCTGGAAGGATCATCTGTTAAACCTTTAAACGCT 252
Qy	200	GAACCTCCAGCAAGGATTCATTTGTTCTCTCGAGGCAACAACCATGTTGCTTCTGCT 259
Db	253	GAACCAAGACAAACGATTTCAATGATTTCTCTCGGCAACAATGGTAGCAAGAAATTGCA 312
Qy	260	TCTGAGACGAAGAG-----GAAGATTTTGAACAGATAGCCAGTGATCTCGACAAT 310
Db	313	GAGGAAGTTGAAGTGGTTGAGATTTGAGATTTTGAAGAGCTTCTAAGAAAGTTAGAGAT 372
Qy	311	GCTTCACCTCTTGAATCATGGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
Db	373	GCTTCACCTCTTGAGATTATGACAAAGCTCTTGAGAAATACGGGAACGATATGCCATT 432
Qy	371	GCTTTCAGTGTGCTGAAGATGTTGCTTTGATTGATATGCGAAATTCACGGTCCAGCC 430
Db	433	GCAITTTAGTGTGAGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGGAGGCCA 492
Qy	431	TTTAGGGTTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTTCAT 490
Db	493	TTTAGAGTATTAGTTTGGATACAGGGAGGTTGAACTCTGAGACGTATCGGTTTTTCGAT 552
Qy	491	GCGGTTGAGAAAGCAATTATGGAATTCGCATTCAGTATCATGTTCCCTGATGCTGTTGAGGTT 550
Db	553	GCGGTGGAGAGCACTATGGGATTAGGATTAGTATATGTTTCTCTGATTTCTGTTGAGGTT 612
Qy	551	CAGGATTGGTGAGGAGTAAGGGTTATTCTTTCTACGAGGATGGGCCCAAGAGTGT 610
Db	613	CAAGGTTTGGTTAGGAGCAAGGGAATGTTCTCTTTTATAGAGATGGTTCATCAGAGAGTGT 672
Qy	611	TGCAGGGTGAGAAAGGTGAGGCTTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGATA 670
Db	673	TGCCGTGTTTCGAAAGGTGAGACCTTTGAGGGTGTCTCTCAAGGGTTTAAAGGCTTGGATT 732
Qy	671	ACTGGTCAGAGGAAGACCACTGCTACTAGGTTCTGAATAACCGGTTGCTTCAGGTT 730

Db 733 ACTGTCAGAGGAAGATCAATCTCCGGGGAAGAGTCTGAGATTCGGTGTTCAGGTT 792
Qy 731 GATCCGGCTTTGAGGGAATGGATGTCGAATTCGAAGCTTGGTGAAGTGAACCCCTGTT 790
Db 793 GATCCGGTGTGTTGAAGGTTGGATGTCGAGTGTGATGTTGGTGAAGTGAATCCCGTT 852
Qy 791 GCAATGTGAAGGCCATGACATATGGAATCTTCCTTAGGACCATGAATGTGCTGTGAAT 850
Db 853 GCGAATGTGAAGGAATGATGTTTGGAACTTCTTAGGACTATGATGTTCCGGTTAAC 912
Qy 851 TCCTTGCATGCAAAAGGATATGTTTCCATTGCGGTGTCAGCCCTGACCTAGGCCCTGTTTA 910
Db 913 ACATTGATCGCGCAGGATATATATGATTCGATTCGAGCTTCGACAAAGCCGTTTA 972
Qy 911 CCTGGCAACATGAAGGGAAGGAGGTGTTGGTGGAGGATGCAAAAGCTTAAGGAATGT 970
Db 973 CCGGTCAGCAGCAGAGAGGAGGAGATGTTGGTGGGAAGATGCTAAAGCCAGGAATGT 1032
Qy 971 GGTCTTCACAAAGGAATGTAAAGCAGCAGAGAGAGGATGTTAATGGAATGGGCTA 1030
Db 1033 GGACTTCACAAAGGAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1089
Qy 1031 TCCCAATCCCATGCAATGGTGTGCTTACCACTGCTGCTGACATTTTCAACAGCCCGAAT 1090
Db 1090 TCGAAT-----CCGCTGTTGCAGATATCTTTAAGAGTGAGAA 1128
Qy 1091 GTAGTTAACTTGACAGCACTGGAATTTGAGAAATTTGCAAAATTTGGAGGACCGAAAGAA 1150
Db 1129 CTTGTGACTTTGACAGCAGGAGGATTTGAGAAATTTGATGAATTTGGAGAACCGTTAAGAG 1188
Qy 1151 CCATGGCTTGTGCTTTATGCAACATGTTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Db 1189 CTTGGATGCTGCTGCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
Qy 1211 TATGTTGACTTAGCAGCAAGTTAGCAGGCTCAACAGGATGAGGTTGGAATTTAGA 1270
Db 1249 TATGATGAATGCGCGGATAAATGGCTGG---AAGTGGGATTAAGGTTGCCAAATTCAGA 1305
Qy 1271 GCAGATGGAGAACAGAAAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACG 1330
Db 1306 GCAGATGCTGACCAAGAGGATTTGCTAAGCAGAAATTCAGCTCGTAGCTTCCCTACC 1365
Qy 1331 ATATTATTTTCCAAAGCAATGCTGCTGCGCAACAATAAGTATCCCTCAGAAAGAGA 1390
Db 1366 ATTCTGTTTTTCCCTAAGAACTATC---GAGACCGATCAAGTATCCGCTCTGAGAGAGA 1422
Qy 1391 GATGTTGATTTCTTGTGCTGCTTGTAAATGCTTAAAGTATGAGGATATCAGGAA 1445
Db 1423 GATGTTGATGCTTTGACTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1477

RESULT 8
ABZ12693
ID ABZ12693 standard; DNA; 1398 BP.
XX AC ABZ12693;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 498.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN W0200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US026685.
XX PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Harper JP, Kreps J, Wang X, Zhu T;

WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 498; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

SQ Sequence 1398 BP; 357 A; 219 C; 407 G; 415 T; 0 U; 0 Other;

Query Match 40.0%; Score 651.6; DB 6; Length 1398;

Best Local Similarity 70.9%; Pred. No. 6.1e-177;

Matches 966; Conservative 0; Mismatches 354; Indels 42; Gaps 6;

Qy 83 TCATCGGAACCCAAATTCGCAAAATTTGGTCAATTAGGATTTGGGAGAGCCCATTTGA 142

Db 67 TCATTTGAGGCAAAAGTTTCGCAAAATTTGGTTCGTGAGGTTATTTGGATCGTTCATGTT 126

Qy 143 GGGCGCGCTTAATTTCAATTTATCTCAAAAGCGG---AGCTTGGTAAAGCCGTTAAAGCC 199

Db 127 GCTCTGTGTCTCGAATCTATCTGGGAAGGATCATCATCTGTAAACCTTTAAACGCT 186

Qy 200 GAACCTCCAGCAAGGATTCATTTCTCTCGAGCAACAACCATCTGTTGCTTCTGCT 259

Db 187 GAACCAAGCAAAAGGATTCATATGATTTCTCTTGGGCAACAATGTTAGCAGAAATGCA 246

Qy 260 TCTGAGACGAAGAG-----GAGATTTTGAACAGATAGCCAGTCTCGACAAT 310

Db 247 GAGGAAGTTGAAGTGGTTGAGATTTGAGATTTTGAAGAGCTTGTCTAAGAGTTAGAGAT 306

Qy 311 GCTTCACCTCTGAAATCATGGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370

Db 307 GCTTCACCTCTGAGATTATGGACAAAGCTTTGAGAAATACGGGAAGGATATCGCCATT 366

Qy 371 GCCTTCAGTGGTCTCGAAGATGTTGCTTGAATGAGTATCGGAAATTTGACGGTTCGACCC 430

Db 367 GCATTTAGTGGTGCAGAAAGATGTTGCTCTTATTTAGTACGCTCATTTGACTGGGAGCCA 426

Qy 431 TTTAGGGTTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTAT 490

Db 427 TTTAGAGTATTTAGTTTGGATACAGGAGGTTGAATCTCTGAGACGATATCGGTTTTTCAT 486

Qy 491 GCGGTTGAGAGCAATTTAGAAATTCGATTTGAGTACATGTTTCCCTGATGCTGTTGAGGTT 550

Db 487 GCGGTCGAGNAGCACTATGGGATTTAGATTAGTATATGTTTCTGATTTCTGTTGAGGTT 546

Qy 551 CAGGCAATGCTGAGGAGTAAAGGGTTATTTCTTTTCTACAGGATGGCCACCAAGAGTGT 610

Db 547 CAAGGTTTGGTTAGGAGCAAGGATTTCTCTTTTTTATGAGGATGCTCATCAGGAGTGT 606

Qy 611 TGCAGGTTGAGAAAGGTGAGGCTTTAAGGAGGCGCCCTTAAGGGTCTCAGAGCATGATA 670

Db 607 TGCCGTTTTCGAAAGGTGAGACCTTTGAGGCGTCTCTCAAGGGTTTTAAAGGGCTTGGATT 666

QY 671 ACTGGTCAGAGGAAGACCCAGTACCTGGTACTAGGTCTGAAATACCGGTTGTTGAGTTT 730
DB 667 ACTGGTCAGAGGAAGAGATCAATCTCCGGGGCAAGGTCAGATTCGGGTTGTTGAGTTT 726
QY 731 GATCCGGCTTTGAGGGAATGGAATGGATGGTGGAGCTTGGTGAAGTGGAAACCCCTGTT 790
DB 727 GATCCGGCTTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCTTT 786
QY 791 GCAATGTGAAGGCCCATGACATATGGAATCTTCTTAGGACCATGATGTCCTGTGAAT 850
DB 787 GCGAATGTGAAGGGAATGATGTTTGGAACTTCTTGAAGACTATGATGTTTCCGGTTAAC 846
QY 851 TCCTTGTGATCAAAAGGATGTTTCCATTTGGGCTGAGCCCTGACCTAGSCCTGTTTGA 910
DB 847 ACATTGTCATGGCGAGGGTATATATCGAATGGATGGATGGATGGATGGATGGATGGAT 906
QY 911 CCTGGGCAACATGAAAGGAGGAGGTGGTGGGAGGATGCGCAAGCTTAAGGAATGT 970
DB 907 CCGGTCAGCAGAGAGAGGAGATGGTGGTGGGAGATGCTAAAGTGAACCAAGGAATGT 966
QY 971 GGTCTTCAAAAGGAATGTAAAGCAGACAGAAAGAGGAGATGTTAATGGAATGGGCTA 1030
DB 967 GGACTTCAAAAGGGAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACCGGGAA 1023
QY 1031 TCCCAATCCCATGCAATGGTGTGCTACCACTGCTGCTGACATTTTCAACAGCCCGAAT 1090
DB 1024 TCGAAT-----CCGCTGTTCAGATATCTTTAAGAGTGAGAAAT 1062
QY 1091 GTAGTTAACTTTGAGCAGACTGGAATTCAGAAATTTGGCAAAATTTGGAGACCGGAAAGGAA 1150
DB 1063 CTTGTGACTTTGAGCAGCAGGAGATGAGAAATTTGATGAATTTGGAGAACCGTAAAGAG 1122
QY 1151 CCATGGCTTGTGTCTTTATGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
DB 1123 CTTGGATCGTGTGCTTTATGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182
QY 1211 TATGTTGACTTAGCAGACAGTATGACAGGCTCAACAGGATGAGGTTGGAATTTAGA 1270
DB 1183 TATGATGAATCGGCGGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
QY 1271 GCAGATGAGAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACG 1330
DB 1240 CGAGATGCTGACAGAGAGGTTGCTTAAGCAGGAAATTCAGCTCGGTAGCTTCCCTACC 1299
QY 1331 ATATTATTTTCCCAAGCATTCGCTCGGCCCAACAATAAGATATCCCTCAGAAAAGAGA 1390
DB 1300 ATTCGTGTTTCCCTTAAGAACTCATC---GAGACCGATCAAGATATCCGCTCAGAAAGAGA 1356
QY 1391 GATGTTGATTCCTTGATGCGCATTTGTAAATGCCTTAAGATGA 1432
DB 1357 GATGTTGATGCTTTGACTGCTTGTGATCTTGTGATCTTGTCCGATAA 1398

RESULT 9
ADG88184
ID ADG88184 standard; cDNA; 1398 BP.
XX
AC ADG88184;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP4-upregulated pathogen infection-related gene #626.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200222675-A2.
XX
PD 21-MAR-2002.
XX

PF 14-SEP-2001; 2001WO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UUNC-) UNIV NORTH CAROLINA.
PA (GLAZ-) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANGL J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
XX Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX WPI; 2002-292409/33.
DR
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
XX
PS Claim 3; SEQ ID NO 626; 605pp; English.
XX
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
XX ADG87557)) whose expression is altered in response to pathogen infection,
XX and to homologues of these genes from other plants or fungi, especially
XX from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
XX cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
XX expression of genes of the invention was upregulated or downregulated in
XX Arabidopsis plants infected with the oomycete Peronospora parasitica,
XX indicating that they play a role in defence mechanisms. The genes of the
XX invention are regulated by RPP7 or RPP8 which act via unconventional
XX signalling cascades, or by the RPP4-dependent pathway. The invention also
XX relates to polypeptides encoded by the pathogen infection-related genes;
XX promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
XX ; expression cassettes, host cells and pathogen-resistant transgenic
XX plants and their progeny comprising a polynucleotide of the invention;
XX and a method of identifying a plant cell infected with a pathogen. The
XX polynucleotide sequences and methods of the invention are useful for
XX identifying plants infected with a pathogen, and for conferring
XX resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
XX nematodes and insects (e.g., aphids). The present sequence represents an
XX Arabidopsis thaliana gene whose expression is altered in response to
XX Peronospora parasitica infection. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1398 BP; 357 A; 219 C; 407 G; 415 T; 0 U; 0 Other;

Query Match 40.0%; Score 651.6; DB 6; Length 1398;
Best Local Similarity 70.9%; Pred. No. 6.1e-177;
Matches 966; Conservative 0; Mismatches 354; Indels 42; Gaps 6;
QY 83 TCATCGGAACCCAAACTTCGCAAAATGGGTCAATTAGATTTCCGAGAGGCCCAATTGA 142
DB 67 TCATTTGGAGCCAAAGATTTTCGCAAAATGGTTGCTTGGATTTATTCGATCGTTCATGTT 126
QY 143 GCGCCCGTTAATTTCAATTTATCTCAAAGACGG---AGCTTGTAAAGCCCGTTAACGCC 199
DB 127 GCTCTGTGTCTCTGAATCTATCTGGGAAGCGATCATCATCTGTGTTAAACCTTTTAAACGCT 186
QY 200 GNACTCCACGCMAGGATTCATTTGTTCTCTCGCAGCAACCAACCATCGTTCCTTCTGCT 259
DB 187 GAACCAAGACCAAGATTTCAATGATTTCTCTTCGGGCAACAAATGGTAGCAGAAATTGA 246
QY 260 TCTGAGACGAAAGAG-----GAAGATTTTGAACAGATAGCCAGTGTATCTCGACAAT 310
DB 247 GAGGAGTTGAAGTGTGTTGAGATTTTGAAGAGCTTGTCTAAGAAAGTTAGAGAAT 306
QY 311 GCTTCACCTTTGAAATCATGATGAGCCCTCGACAAATTCGGCAACGACATAGTATT 370
DB 307 GCTTCACCTTTGAGATTATGCAAAAGCTCTTTGAAATATCGGGAACGATATCGCCATT 366

Qy	371	GCCTTCAGTCGGTGCTGAAAGATGTGCTTTGATTTGAGTATGCGAAATATGACGGGTGACCC	430
Db	367	GCATTTAGTGGTGACAGAAATGTGTCTTATTTGAGTACGCTCATTTGACTCGGAGGCCA	426
Qy	431	TTTATAGGGTTTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTTCGAT	490
Db	427	TTTATAGATATTTTATTTTTGGATACAGGAGGTTGAAATCCTTGAGACGTATCGTTTTTCGAT	486
Qy	491	CGCGTTTGAGAAGCAATATATGGAATTCGCATTTGAGTACATGTTCCCTGATCCTGTTGAGGTT	550
Db	487	CGCGTGAGAAGCACATATGGCATATAGGATTTGAGTATATGTTTCTGATTTCTGTTGAGGTT	546
Qy	551	CAGGCATTTGCTGAGGAGTAAAGGGTTTATCTCTTTCTACGAGATCGGCACCAAGAGTGT	610
Db	547	CNAGGTTTGGTTATAGGACCAAGGATTTGTTCTCTTTTATAGAGTATGTTATCAGGAGTGT	606
Qy	611	TGCAGGCTGAGAAAGGTGAGCCCTTTAAAGAGAGGCCCTTTAAAGGCTCTCAGAGCATGGATA	670
Db	607	TGCGTGTTCGAAAGGTGAGACCTTTGAGGCGTCTCAAGGGTTTAAAGGCTTCGATTT	666
Qy	671	ACTGGTCAGAGGAAAGACAGTCACTTGGTACCTAGGTCTGAAATACCGGTTGTTTTCAGGTT	730
Db	667	ACTGGTCAGAGGAAAGATCAATCTCCGGGACAAGGCTCTGAGATTCGGGTTGTTTTCAGGTT	726
Qy	731	GATCCGGCTTTTGAGGGAAATGGAATGTTGGAAGCTTTGGTGAAGTGGAAACCCCTGTT	790
Db	727	GATCCGGTGTTTGAAGTTTGGATGTTGGAGTTGTTGTTGTTGGTGAAGTGGATCCGGTT	786
Qy	791	GCAAATGTGAAGGCCCATGATCATATGGAACCTTCCTTAGGACCAATGAATGTGCTGTGAAAT	850
Db	787	GCGAATGTTGAAGGGAAATGATGTTTGAACCTTCTTGAGGACTATGGATGTTCCGGTTAAC	846
Qy	851	TCCTTGATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTTTA	910
Db	847	ACATTTGATCGGCAGGGGTATATATCATTTGGAATGTGAGCTTGTGCACGAAAGCGGTTTTA	906
Qy	911	CCTGGGCAACATGAAAGGGAAGGAGGTGCTGTGTGGGAGGATGCCAAGACTAAGGAAATGT	970
Db	907	CCGGGTGACACGACGAGAAAGGAGATGTTGGTGGGAGATGCTTAAGCCAGGAATGT	966
Qy	971	GGTCTTCAAAAGGAAATGTAAAGACGACGAGAAAGAGGAGATGTTAATGAAATGGGCTA	1030
Db	967	GGACTTCAAAAGGGAATGTCA---AAGAAAACTCCGATGCTAAAGTGAACGGGGAA	1023
Qy	1031	TCCCAATCCCATGCAATGGTGTGCTACCACTGTGCTGCATGCTTTCACACTTTTCAACGCCGAAT	1090
Db	1024	TCGAAT-----CGCTGTTGCAGATATCTTTAAGATGAGAAAT	1062
Qy	1091	GTAGTTAACTTTGACGAGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAA	1150
Db	1063	CTTGTGACTTTGACGAGCGAGGGGATTTGAGAAATTTGATGAAGTTGGAGAACCGTAAGAG	1122
Qy	1151	CCATGGCTTGTGTGCTTTATGACCAATGTTGCCCTTACTCCAGGCGTATGGAGGAATCT	1210
Db	1123	CCTTGGATGCTGCTGCTTTATGCTCCGTTGGTGGCCCCCTTTGTCAAGCCATGGAGATCG	1182
Qy	1211	TATGTTGACTTTAGCAGACAAGTTAGCAGGGGTCAACAGGGATGAAGTTGGAAATTTAGA	1270
Db	1183	TATGATGAACCTGSCGGATAAATTCGCTGG---AAGTCGGATTAAGGTTGCCAAATTCAGA	1239
Qy	1271	GCAGATGAGAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAAGTTCCTTACG	1330
Db	1240	GCAGATGGTGAACCAAGAGGAGTTTGCTAAGCAGGAATTCGACGTTCGGTATCTCCCTTACC	1399
Qy	1331	ATATTTATTTTTCCCAAAGCAATTGCTCTCGGCCCAACAATAAGTATCCCTCAGAAAAAGAGA	1390
Db	1300	ATTCTGGTTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGAGAGAGA	1356
Qy	1391	GATGTTGATTCCTTGATGGCAATTTGTTAAATGCCCTTAAGATGA	1432
Db	1357	GATGTTGAGTCTTTGACTTCGTTCTTTGAAATCTTCTGCGATAA	1398

[illegible]

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-014819P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.1%; Score 588.2; DB 3; Length 1687;
Best Local Similarity 69.4%; Pred. No. 1.3e-158;
Matches 895; Conservative 0; Mismatches 343; Indels 51; Gaps 5;
QY 157 CAATTTATCTCAAGACGGAGCTTGTTAAAGCCCGTTAAACGCCGA-----ACCTCCACG 210

Db 217 CCATTGCTCAGAGAGCTTACTCTATGAAACCCCTTAAACGCTGAGTCATCACATACAAG 276
Qy 211 CAAGGATTCATGTTCTCTCGCAGCAACAACCATCGTTGCTTCTCTCTCTGAGA--- 266
Db 277 AAGCGATCTGGGTTCTTCGTGCTTCACTCTAATTCCTCTGCTGAGTGAAGAGAAGA 336
Qy 267 -CGAAGAGAGAGATTTGAAACAGATAGCAGCAGTGTCTCGACAATGCTTCACCTTTGA 324
Db 337 AGGAGAAGTTCAAGACTTTGAGCAACTTGTCTAAAGCTTTGAAGATGCTTCTCCACTTGA 396
Qy 325 AATCATGATAGAGCCCTCGACAAATTCGCAAGCATAGCTATTGCTTCAGTGTGTC 384
Db 397 AATCATGGATAAAGCTCTTGAGAAATTTGAGAGCGAAATCGCAATTCCTTTTATGAGGAGC 456
Qy 385 TGAAGATGTTGCTTTGATTCAGTATGCGAAATTAACCGGTCGACCCCTTTAGGGTTTTTCAG 444
Db 457 TGAAGATGTTGCATTGATGATGATATGCTGCTTAACTGGAAGCCATTTAGGGTTTTTAG 516
Qy 445 TTTGGACACTGGGAGACTGAACCCAGAAATTTATCAACTTTTGTGATCGGTTGAGAAGCA 504
Db 517 TTTAGATACAGGGAGATTAACCCGGAACGTCACAGGCTCTTTGACGCGAGTCGAGAAGCA 576
Qy 505 TTATGGAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGCATTTGTGAG 564
Db 577 GTACGGGATTCGAAATTCGATGATGTTTCTGATGCGAGTTGAGGTTTCAAGCTTTAGTGAG 636
Qy 565 GAGTAAGGGGTTATTCCTTTCTACGAGGATGGCCACCAAGAGTGTTCAGGGTGTGAGAA 624
Db 637 GAACAGGGGTTGTTCTATCTTATGAAGATGTCATCAAGAGTGTTCGCGTGTGAGGAA 696
Qy 625 GGTGAGCCCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGATTAACCTGGTCAGAGGAA 684
Db 697 AGTTAGACCTTTGCGTCTGCTCTTAAGGGTCTTAAAGCTTGGATTACAGGACAGAGGAA 756
Qy 685 AGACCACTCCTGGTACTAGTCTGAATACCGGTTGTTTCAGGTTGATCCGCTTTTGA 744
Db 757 AGACCAATCTCCGGTACGAGATCTGAGATCCCTATTGTTTCAGGTTGATCCAGTGTGTA 816
Qy 745 GCGAATGGATGGTGGAAATTCGAAGCTTGGTGAATGGAGAACCCCTGTTGCAAAATGTGAAGG 804
Db 817 AGGGTTAGATGGCGGTGTGGAAGTCTTGTGAAGTGAATCCCTTTGGCTAATGTTGAAGG 876
Qy 805 CCATGACATATGGAATCTCTTAGGACCATGAATGTGCTGTGAATTCCTTGTGATGCATA 864
Db 877 AGCTGATGTGTGAACCTTTCTGAGAACTATGGATGTTCCGGTGAATGCAATGTCATGCACA 936
Qy 865 AGGATATGTTTCCATTTGGGTGTGAGCCCTGCTAGCCCTGTTTACCTGGGCAACATGA 924
Db 937 AGGATATGTGTCAATCGGGTGTGAGCCGTGTACTAGGCCCTGTGCTTCAGGTCAGCATGA 996
Qy 925 AAGGGAAGGAGGTGGTGGGAGGATGCCAAGCTTAAGGAATGTGCTTTCACAAAGG 984
Db 997 AAGAGAAGGAAGGTGGTGGGGAAGATGCTAAAGCTTAAAGAAATGTGCTTACACAAAGG 1056
Qy 985 AAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGC 1044
Db 1057 GAACATCAAGGAGGAAGTGTG-----CTGC 1083
Qy 1045 AAATGGTGTGCTACCACTGTGCTGCTGATTTTCAACAGCCCGAATGTAGTTAACTTTGAG 1104
Db 1084 AGACTCAAGCCCTGCTGCTGCAAGAGATATTTGAAGCAACAATGTGTTGTCATTGAG 1143
Qy 1105 CAGSACTGGAATTTGGAATTTGCAAAATTTGGAGGACCGAAGGAACCATGGCTTGTGT 1164
Db 1144 CAGAGAGGGGTTGAGAAATCTTTTGAAGCTAGAGAACCGTAAAGAGCGGTGGCTGTGCT 1203
Qy 1165 GCTTTATGCAACCATGTGTCCTTACTGCCAGGCTATGAGGGAATCTTATGTTGACTTAGC 1224
Db 1204 ACTTTAGCTCTTCTGTCGCTTCTGTCAGGCGATGGAAGCATCGTACGTCGAATTTGCG 1263
Qy 1225 AGACAAGTTAGCAGGGTCAACAGGGAATGAAGGTTGGAATTTTAGAGCAGATGGAGAAC 1284

Db 1264 TGAGAACTTGGCGG---AAAAGGAGTTAAAGTGGCGAAATTCGAGCTGACGGTGAGCA 1320
Qy 1285 GAAAGAAATTTGCAAGAGTGAACTGCAATTTGGGAAGCTTCCCTACGATATTTTTCCTCC 1344
Db 1321 GAAGGATTTGCTTAAGCAAGAGCTTCAGTTAGGAGCTTCCGACGATACTTCTCTTCC 1380
Qy 1345 AAAGCATTCGCTCGGCGCAACAATAAGATATCCCTCAGAAAAGAGAGATGTTGATTCCTT 1404
Db 1381 GAAAGAGCTCCACGG---GCTATTAAAGTACCCTTCAGAGCATAGAGATGTTGATCACT 1437
Qy 1405 GATGGCAATTTGTAATCCCTTAAAGATGAG 1433
Db 1438 CATGCTGTTGTAATCTTCTTCGGTGAG 1466

RESULT 13
AAC48356
ID AAC48356 standard; DNA; 1685 BP.
XX
AC AAC48356;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57179.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

35.5%; Score 578.4; DB 3; Length 1685;

	Best Local Similarity	69.0%;	Pred. No. 9e-156;	Matches 887;	Conservative	0;	Mismatches 351;	Indels	48;	Gaps	5;
Qy	157	CAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACGCCGA---ACCTCCACGCCAA	213								
Db	219	CCATTTGTCTCAGAGACGGTACTCTATGAACACCTTTAAACCGTGAGTCACATTCACGAG	278								
Qy	214	GGATTCATTTGTTCTCTCAGAGCAACAAACCATCGTTCCTCTGCTCTCTGAGA-----C	267								
Db	279	CGAATCTTTGGGTAACTCGTCTCTCAACTCTAATTTGCTCTGAACTTGAAGAGAAAGGAGG	338								
Qy	268	GAAGAGGAGACATTTTGAACAGATAGCCAGTGATCTCGACAACTGCTTCACCTCTTTGAAT	327								
Db	339	AGAAGTTGAAGACCTTTGAGCACTTGCTTAAAGAGCTTGAAGATGCTTCTCCACCTTGAAT	398								
Qy	328	CATGGATAGAGCCCTCGACAAATTCGGCAACCGACATAGCTATTGCCCTTCAGTGTGCTGA	387								
Db	399	CATGGATAAAGCTCTTGAGAGATTCGGAGACCAATCGCAATTCCTTTAGTGGAGCTGA	458								
Qy	388	AGATGTTGCTTTGATTTGAGTANGCAAAATTCAGCGGTGACCCCTTTAGGGTTTTACGTTT	447								
Db	459	AGATGTTGCAATTTGATTTGAATATGACACGTTTAACTGGAAGGCCATTTAGGGTTTTAGT	518								
Qy	448	GGACATCGGGAGACTGCAACCCAGAACTTATCAACTTTTTCATCGCGTTCAGACGCTTGA	507								
Db	519	AGATACAGGGAGATTTAAACCCCTGAAACGTTACAGGCTCTTTGACCGAGTCGAGAGCAG	578								
Qy	508	TGGAATTCGCATTCAGTACATGTTTCCCTGATGCTTGTGAGGTTTCAGAGCATTTGGTAGG	567								
Db	579	CGGATTTGCAATTCAGTACATGTTTCTGATCGAGTTTTCAGGTTTCAGGTTTTCAGGAG	638								
Qy	568	TAAGGGTTTATCTCTTTCTACAGGATGGGACCAAGAGTGTGTCAGGGTTCAGGAGAAAG	627								
Db	639	CNAGGGTTTGTCTCTCAATTCATGAAGATGGTTCATCAAGAGTGTTCGCGTGTGAGGAA	698								
Qy	628	GAGGCCCTTTAAGAGGGCCCTTAAAGGCTCTCAGAGCATGGAATACCTGTCAGAGGAA	687								
Db	699	TAGACCTTTGCGTCTGCTCTTAAAGGGTCTTAAAGCTTGGATTTACAGGACAGAGGAA	758								
Qy	688	CCAGTCACTGGTACTAGTCTGAAATACCGGTTGTTTCAGTTCATCCGCTTTTGAGGG	747								
Db	759	CCAATCTCCGGGTACGAGATCTGAGATCCCTATTTGTTTCAGTGTGATCCAGTGTGAG	818								
Qy	748	AATGGATGGTGGAAATGGAGCTTTGGTGAAGTGGAAACCTGTTGCAATGTGAAGGGCCA	807								
Db	819	GTTAGATGGCGGTGTTGGAAGTCTTGTAAGTGGAAATCCCTTTCGCTAATGTTGAAGG	878								
Qy	808	TGACATATGAACTTCTCTTAGGACCATGAATGTGCTCTGTAATTCCTTTGATGCAAAAG	867								
Db	879	TGATGTCGGAACCTTCTGAGAACTATGAGATGTTCCGGTGAATGCAATGCAACGCAAG	938								
Qy	868	ATATGTTTCCATTTGGGTGTGAGCCCTGCACCTTAGSCCTGTTTTCCTGGGCAACATGAAG	927								
Db	939	GTAATGTCTCAATCGGGTGTGAGCCGTGTACTAGCCCGGTGCTTCAGGCCAAATGAGAG	998								
Qy	928	GGAAGGAGGTGTGGTGGGAGGATGCCAAAGCTTAAGGAATGTGGTCTTCACAAAGGAAA	987								
Db	999	AGAAGGAAGGTGTGGTGGGAGATGCTAAAGCTTAAGATGTGGTCTACACAAAGGAA	1058								
Qy	988	TGTAAGACGACGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCCAAA	1047								
Db	1059	CATCAAGGAGGAAGATGGTG-----CTGCAGA	1085								
Qy	1048	TGGTGTGCTACCACTGTGCTCTGACATTTTCAACGCCCGGAATGTAGTTAACTTGAGCAG	1107								
Db	1086	CTCAAGGCCCTGTGCTGTGCAAGAGATATTTGAAGCAACAATGTGGTGTGCAATGAGCAA	1145								
Qy	1108	GACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGGAAAGGAAACCATGCTGTGTG	1167								
Db	1146	AGGAGGGTTGAGAACTTTTGAAGCTAGAGAACCGTTAAGAGCGGTGGTGTGCTACT	1205								
Qy	1168	TTATGCAACCATGTGCCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGCACTTAGCAGA	1227								

Db	1206	TTACGCTCCTTGGTGGCCCTTTCTGCCACGGCGATGGAGCATCGTACATCGAAATTGGCTGA	1266
Qy	1228	CAAGTTAGCAGGGTCAACACGGGATGAAGTTGGAAATTTAGAGAGCAGATCGGAGAACAGAA	1287
Db	1266	GAACTTTGCGGG---AAAGAGGTTAAAGTGGCGAAATCCGAGCTGACGGTGGCAGAA	1322
Qy	1288	AGAAATTTGCAAAAGATGAACTGCAATTTGGGAAGCTTCCCTPACGATATTAATTTTCCCAA	1347
Db	1323	GGAGTTTGTCTAAAGCAAGAGCTTCAGTTAGGAGCTTCCCGACGATACTTCTCTTTCCGAA	1382
Qy	1348	GCATTCGCTCGCCCAACATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGAT	1407
Db	1383	AAGAGCTCCACGG---GCTATTAGTACCCTTCAGAGCATAGAGATGTTGATTCACCTCAT	1439
Qy	1408	GGCATTTTGTAATGCCCTTAAGATGAG	1433
Db	1440	GTCGTTGTGATCTTCTTCGGTGAG	1465
RESULT 14			
ID	ABZ14500	standard; DNA; 1365 BP.	
XX	AC	ABZ14500;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2305.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PP	24-AUG-2001; 2001WO-US026685.	
XX	PR	24-AUG-2000; 2000US-0227866P.	
XX	PR	26-JAN-2001; 2001US-0264647P.	
XX	PR	22-JUN-2001; 2001US-0300111P.	
XX	PA	(SCRI) SCRIPPS RES INST.	
XX	PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Harper JF, Kreps J, Wang X, Zhu T;	
XX	DR	WPI; 2002-304127/34.	
XX	PT	Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.	
XX	PS	Claim 144; SEQ ID NO 2305; 577pp + Sequence Listing; English.	
XX	CC	The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office	
XX	SQ	Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;	
Query Match			
Best Local Similarity 68.4%; Score 577.4; DB 6; Length 1365;			
Matches 886; Conservative 0; Mismatches 351; Indels 48; Gaps 5;			

QY 157 CAATTATCTCAAGACGGAGCTTGGTAAGCCGGTTAAACCCGA---ACCTCCACGCAA 213
Db 120 CCATTTGTCTCAGAGACGTTACTCTATGAACCTTTAAACCGCTGAGTCAATTCACGAAG 179
QY 214 GGATTCATTTGCTCTCCGACGACAAACATCGTTGCTTCTGCTTCTCTGAGA-----C 267
Db 180 CGAATCTTGGGTACTCGTGTCTCACTCTAAATGCTCTCTGAAAGTTGAAGAAAGGAGG 239
QY 268 GAAAGAGGAAGATTTTGAAACAGATAGCAGTGATCTCGACAATGCTTCACTCTCTTGAAAT 327
Db 240 AGAAGTTGAAGACTTTGAGCAACTTGCTAAAGAGCTTTGAAGATGCTTCTCCACTTGAAT 299
QY 328 CATGATAGACCTTCACAAATTCGCAACGACATAGCTATTGCTTCACTGATGCTGCTGA 387
Db 300 CATGGATAAAGCTCTTCAGAGATTCGAGACCAATCGCAATTCCTTTAGTGAGCTGA 359
QY 388 AGATGTTGCTTTGATTCAGTATGCAAAATTCACGGGTTCGACCTTTAGGGTTTTTCAGTTT 447
Db 360 AGATGTTGCAATGATGATATGACCGTTTAACTGGAAGCCATTTAGSGTTTTAGTTT 419
QY 448 GGCACCTGGGAGACTGAACCCAGAAACTTATCAACTTTTTGATGCGGTGAGAAAGCATTA 507
Db 420 AGATACAGGGAGATTAAACCTCGAAACGTAACGGCTCTTTGACGCACTCGAGAAGCAGTA 479
QY 508 TGGAAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGGTGAGAG 567
Db 480 CGGGATTCGAAATTCAGTACATGTTTCTGATGCAAGTTTGAAGTTTCAAGCTTTAGTGAGAA 539
QY 568 TAAGGGTTATTCCTTTCTACGAGGATGGCAACAGAGTGTTCGAGGGTGAGAAAGGT 627
Db 540 CAAGGGTTTGTCTCTATCTATGAAGATGTCATCAAGAGTTGCGGTGAGGAAAGT 599
QY 628 GAGGCCCTTTAAGGAGGCGCTTAAGGGTCTCAGAGCATGGAATACTGCTCAGAGGAAGA 687
Db 600 TAGACCTTTGCGTGTCTCTTAAGGGTCTTTAAGGCTTGAATTAACGACAGAGCAAGA 659
QY 688 CCACTGACCTGGTACTAGTCTGAATACCGGTTGTTGAGGTTGATCCGGCTTTTGGGG 747
Db 660 CCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCAGGTTGATCCAGTGTGTAAGG 719
QY 748 AATGATGTTGGAATTCGAAGCTTGGTGAAGTGAACCTCTGTCAAATGTGAAGGGCCA 807
Db 720 GTTAGATGCGCGTGTGGAAGCTTTGGAAGTGGATCCCTTTGGCTAATGTTGAGGAGC 779
QY 808 TGACATATGGAACCTTCCTTAGGACCATGAATGTGCTGTGAAATTCCTTTGCAATGCAAAAGG 867
Db 780 TGATGTGTGGAACCTTCTGAGAACTATGGAATGTTCCGGTGAATGCAITTCACGCGCACAGG 839
QY 868 ATATGTTTCCATTCGGTGTGAGCCCTGCATGAGCCTGTTTACTCTGGCAACATGAAG 927
Db 840 GTATGTGTCAATCCGGTGTGAGCCGTGTACTAGGCGGTGCTTCAGGCCCAACATGAGAG 899
QY 928 GGAAGGAGGTGTTGTTGGGAGGATGCCAAAGCTAAGGAATGTGCTCTTCACAAAGGAAA 987
Db 900 AGAAGGAAGTGTGTTGGGAGAGATGCTTAAAGCTAAAGAAATGTGCTTACACAAAGGAA 959
QY 988 TGTAAAGCAGCAAGAAAGAGGAGTGTTAATGGAATGGGCTATCCCAATCCCATGCAAA 1047
Db 960 CATCAAGGAGGAAGATGTG-----CTGCAGA 986
QY 1048 TGGTGATGCTACCACTGTGCTGTGACATTTTCAAGAGCCCGAATGTAGTTAACTTTGAGCAG 1107
Db 987 CTCGAAGCCTGTGCTGTGCAAGAGATATTTTGAAAGCAACAAATGTTGGTTGTCATGAGCAA 1046
QY 1108 GACTGGAAATTCAGAAATTTGGCAAAATTCGAGGACCGAAGGAACCATGGCTTGTGTGCT 1167
Db 1047 AGGAGGGTTGAGAACTTTTGAAGCTAGAGAACCGTAAAGAGGCGTGTGTTGCTGCTACT 1106
QY 1168 TTATGCAACATGTTGCCCTTACTGCTCCAGGCTATGGAAGGAATCTTATGTGACTTTAGCAGA 1227
Db 1107 TTACGCTCCTTGGTCCCTTTCTGCCGAGCGATGGAAGCATCGTACATCGAATTCGCTGTA 1166

QY 1228 CAACTTAGCAGGTCAACAGGGATGAAGTTGGAAATTTTGAAGCAGATGAGAAACAGAA 1287
Db 1167 GAACTTGGGG---AAAGGAGTTAAAGTGGCGAAATTCGAGCTGACGGTGAGCAGAA 1223
QY 1288 AGAATTTGCAAGAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATATTTTCCCAA 1347
Db 1224 GGAGTTTGTCTAAGCAAGAGCTTCAGTTAGGGAGCTTCCCGACGATACTTCTCTTCCGAA 1283
QY 1348 GCATTCGTCTCGCCCAACATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGTAT 1407
Db 1284 AGAGCTTCCAGG---GCTATTAGTACCTTCAGACATAGAGATGTTGATTCACATCAT 1340
QY 1408 GGCATTTGTAAATGCCTTAAAGATGA 1432
Db 1341 GTCGTTGTGAATCTTCTTCGGTGA 1365
RESULT 15
ADG88214
ID ADG88214 standard; cDNA; 1365 BP.
XX ADG88214;
AC ADG88214;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP4-upregulated pathogen infection-related gene #656.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
FN WO200222675-A2.
PN
PD 21-MAR-2002.
XX
PP 14-SEP-2001; 2001WO-US028506.
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANG J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX WPI; 2002-292409/33.
DR
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
PT
XX
PS Claim 3; SEQ ID NO 656; 605pp; English.
XX
CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557)) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;

CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

Query Match 35.4%; Score 577.4; DB 6; Length 1365;

Best Local Similarity 68.9%; Pred. No. 1.6e-155;

Matches 886; Conservative 0; Mismatches 351; Indels 48; Gaps 5;

Qy	157	CAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACCCGA---ACCTCACGCA	213
Db	120	CCATTTGTCTCAGAGACGTTACTCTATGAACCTTTAAACGCTGAGTCACATTCACGAAG	179
Qy	214	GGATTCCATTCTCTCGCAGCAACCAACATCGTTGCTTCTGCTCTGAGA-----C	267
Db	180	CGAATCTTGGGTACTCTGCTTCACTCTAATTGCTCTCGAAGTTGAAGAGAAAGGAGG	239
Qy	268	GAAGAGAGAGATTTTGAACAGATAGCCAGTGATCTCGACAAATGCTTCACCTCTTGAAT	327
Db	240	AGAAAGTTGAAGACTTTGAGCAACTTGTAAAGCTTGAAGATGCTTCCACTTGAAT	299
Qy	328	CATGGATAGACCTCTCGACAAATTCGGCAACGACATAGCTAATTCCTTCAGTGTGCTGA	387
Db	300	CATGGATAAGACTCTTGAGAGATTCGGAGACCAAAATCGCAATTCCTTTAGTGAGCTGA	359
Qy	388	AGATGTTGCTTTGATTCAGATATGCAAAATTCACCGGTGACCTTTAGGGTTTTCACTTT	447
Db	360	AGATGTTGATTTGAATATGACGTTTAACTTGAAGAGCCATTTAGGGTTTTTAGTTT	419
Qy	448	GGACACTGGGAGACTGAACCCAGAAACTTTATCAACTTTTGTATCGGTTGAGAAAGCATTA	507
Db	420	AGATACAGGGAGATTAAACCCGTGAACGTCAGAGCTTTTGACGAGTCGAGAGACGATA	479
Qy	508	TGGAATTCGCAATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCAAGGATTCAGGAG	567
Db	480	CGGATTCGAATTCAGTACATGTTTCTGATGACGATTCAGGTTTCAAGCTTTAGTGAGGA	539
Qy	568	TAAGGGTTATCTCTTTCTACGAGGATGGCCACAGAGTGTGACGGTGTGAGAGGT	627
Db	540	CAAGGGTTTGTCTCATTTCTATGAAGATGGTTCATCAAGAGTGTGCGGTGTGAGGAAAT	599
Qy	628	GAGGCTTTTAAAGAGGGCCCTTAAAGGGTCTCAGAGCATGGATACTGGTCAGAGGAAGA	687
Db	600	TAGACCTTTGGTCTGCTCTTAAAGGGTCTTAAAGCTTGGATACAGACAGAGGAAGA	659
Qy	688	CCAGTCACTGGTACTAGGTCTGAATAACCGGTTGTTCAAGTTGATCCGGCTTTTGAAGG	747
Db	660	CCAACTCTCCGGGTACGAGATCTGAGATCCCTTATTTGTTTCAAGTTGATCCAGTGTGAAG	719
Qy	748	AATGGATGGTGAATTCGAACTTGGTGAAGTGAACCCCTGTTGCAATGTGAAGGCCA	807
Db	720	GTTAGATGGCGGTGTGGAAGTCTTGTGAAGTGAATCCCTTTGGCTAATGTGAAGGAGC	779
Qy	808	TGACATATGGAACTTCCCTTAGGACCATGAATGTGCTGTGAATTCCTTGTGATGCAAAAGG	867
Db	780	TGATGTCTGGNACTTTCTGAGAACTATGGATGTTCCGGTGAATGCAATGCAACGCAAGG	839
Qy	868	ATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGAAG	927
Db	840	GTATGTCTCAATCGGGTGTGAGCCGTGCTACTAGGCCGCTGCTCCAGGCCCAACATGAGAG	899
Qy	928	GGAGGAGGTGTTGGTGGGAGGATCCAAAGCTTAAGGAATGTGCTTTCACAAAGGAA	987
Db	900	AGAAGGAAGGTGGTGGGAGAGATGCTAAAGCTTAAAGATGTGGTCTCACAAAGGGNA	959

Search completed: February 20, 2006, 21:00:54
Job time : 951 secs

Qy	988	TGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAA	1047
Db	960	CATCAAGAGGAGAGATGGTG-----CTGCAGA	986
Qy	1048	TGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAG	1107
Db	987	CTCAAGCCTGCTGCTGTGCAAGAGATATTTGAAAGCAACAATGTGGTTGCATTGAGCAA	1046
Qy	1108	GACTGGAAATTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGCAACCATGGCTTGTGTGCT	1167
Db	1047	AGGAGGGGTTGAGAAATCTTTTGAAGCTAGAGAACCGTAAAGAGGGGCTGTTGGTCTACT	1106
Qy	1168	TTATGCCACCATGGTGGCCCTACTCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGA	1227
Db	1107	TTACGCTCCTTTGGTGGCCCTTTCTGCCAGGCGATGGAAGCATCGTACATCGAATTTGGCTGA	1166
Qy	1228	CAAGTTAGCAGGGTCAACAGGGATGAAGGTTGGGAAATTTTAGAGCAGATGGAGAACAGAA	1287
Db	1167	GAACCTTTGCGGG---AAAAGGAGTTAAAGTGGCGAAATTTCCGAGCTGACGGTGAAGCA	1223
Qy	1288	AGAAATTTGCAAGAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTTTTCCTCCAA	1347
Db	1224	GGAGTTTGCTAAGCAAGAGCTTTCAGTTAGGGAGCTTCCCGACGATCTTCTCTTTCCGAA	1283
Qy	1348	GCATTCGTCTCGGCAACAATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTCAT	1407
Db	1284	AAGAGCTCCAGG---GCTATTAAAGTACCTTTACAGACATAGAGATGTTGATTCACATCAT	1340
Qy	1408	GGCAATTTGTAAATGSCCTTAAGATGA	1432
Db	1341	GTCGTTTGTGAATCTTCTTCGGTGA	1365

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:45:11 ; Search time 5977 Seconds
(without alignments)

12751.574 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacggaggagaccata.....taaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gss1:*

10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	600.8	36.9	765	2	BE658477 GM700006A
2	599.6	36.8	612	2	BE658404
3	570.6	35.0	1535	4	EX813913 Arabidops
4	567.2	34.8	1579	4	BN504000 Arabidops
5	542.2	33.3	608	2	BF070490 st20h11.Y
6	539.4	33.1	552	2	BF071136 st12f06.Y
7	510.2	31.3	516	3	BI316361
8	504	30.9	830	6	CB339330
9	491.6	30.2	507	3	BI316488
10	486.6	29.9	959	7	CK264757
11	481.4	29.6	629	1	AW508411
12	474.8	29.1	704	6	CB340603
13	473.4	29.1	910	6	CD573900
14	450.8	27.7	764	6	CF119175
15	450.8	27.7	764	6	CF119189
16	446.2	27.4	968	7	CK263241
17	438.4	26.9	716	7	CO492657
18	431.8	26.5	875	6	CB292552
19	429.6	26.4	683	6	CA924320
20	429.2	26.3	660	6	CD487608
21	428.4	26.3	654	2	BG455000
22	427	26.2	1299	10	CL972938

23	426.2	26.2	676	6	CA932068
24	426	26.2	858	2	BF253965
25	425	26.1	856	7	CK287268
26	420.2	25.8	836	9	BH652825
27	420.2	25.8	859	8	DN982440
28	415.6	25.5	963	7	CK266388
29	414.6	25.5	780	3	BM111117
30	413.6	25.4	689	7	CO105768
31	413.2	25.4	671	7	CO105768
32	412.6	25.3	878	6	CF213409
33	412.4	25.3	729	6	CA808814
34	411.4	25.3	831	8	CO534000
35	411	25.2	608	8	CK172578
36	408.4	25.1	689	2	BG439877
37	408.2	25.1	606	6	CD713621
38	408.2	25.1	606	6	CD720052
39	407	25.0	653	2	BG124326
40	406.8	25.0	610	1	AJ558767
41	406.8	25.0	648	6	CD716959
42	406.8	25.0	769	2	BG581920
43	406.2	24.9	729	7	CV100088
44	404.2	24.8	756	10	DU052684
45	401.4	24.6	831	10	CW799574

ALIGNMENTS

RESULT 1
BE658477/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE658477 765 bp mRNA linear EST 24-MAY-2001
GM700006A20E1 Gm-r1070 Glycine max cDNA clone Gm-r1070-2114 3',
mRNA sequence.

BE658477.1 GI:9984369

EST.

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 765)

Vodkin, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
Expanding, J., Raph, C., Shoop, E., Pardini, J., Liu, L., and Lewin, H.
A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AM133461 corresponding to Gm-cl015-1660 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome

systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

source

1. .765

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="Gm-r1070-2114"

/notes="The library Gm-r1070 is a sequence-driven, reracked

set of 9,216 clones selected from cDNA libraries from

various tissues and stages of development of soybean that

represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1070. The cDNA clones of the reracked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Ritzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html>. Reracking was performed by Genome Systems, St. Louis, <http://www.genomesystems.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.life.uiuc.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 36.9%; Score 600.8; DB 2; Length 765;
Best Local Similarity 96.2%; Pred. No. 4.9e-155;
Matches 602; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 988 TGTAAAGCAGCAGAAAGAGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGCAAA 1047
DB 765 TGTNNNNAGCAGANNAGGANGATGTTAATNGANNNGNCTACNNNTCCCATGCAA 706

QY 1048 TGGTGATGCTACCTGTGCTGTGATTTTCAACAGCCCGAATAGTTAACTTTGAGCAG 1107
DB 705 TGGTGATGCTACCTGTGNNNGACAATTTCAACAGCCCGAATAGTTAACTTTGAGCAG 646

QY 1108 GACTGGAAATTCAGAAATTTGGAGGATCGGAGCCGAAAGCAATGGCTTTGTTGCT 1167
DB 645 GACTGGAAATTCAGAAATTTGGAGGATCGGAGCCGAAAGCAATGGCTTTGTTGCT 586

QY 1168 TTATGCACCATGGTGCCCTACTGCCAGGCTATGGAGGAATCTTATGTGACTTTAGCAGA 1227
DB 585 TTATGCACCATGGTGCCCTACTGCCAGGCTATGGAGGAATCTTATGTGACTTTAGCAGA 526

QY 1228 CAAGTTAGCAGGGTCAACAGCGATGAAGTTGGAAAATTTAGACAGATGAGAACAGAA 1287
DB 525 CAAGTTAGCAGGGTCAACAGCGATGAAGTTGGAAAATTTAGACAGATGAGAACAGAA 466

QY 1288 AGAATTTGCAAGAGTGAAGTGAATTTGGGAAGCTTCCCTACGATATATTTTCCCAAA 1347
DB 465 AGAATTTGCAAGAGTGAAGTGAATTTGGGAAGCTTCCCTACGATATATTTTCCCAAA 406

QY 1348 GCATTCGTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGATGTTGATTCCTTGAT 1407
DB 405 GCATTCGTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGATGTTGATTCCTTGAT 346

QY 1408 GGCATTTTGAATGCTTTAAGATGAGGATATCAGGAAAATTTCTTCGTTTTGGGTGCA 1467
DB 345 GGCATTTTGAATGCTTTAAGATGAGGATATCAGGAAAATTTCTTCGTTTTGGGTGCA 286

QY 1468 ATTCACATTTGACTATAGTACAGCGGTTCTTCTTTATGCTATTAGTGTATATACCA 1527
DB 285 ATTCACATTTGACTATAGTACAGCGGTTCTTCTTTATGCTATTAGTGTATATACCA 226

QY 1528 TTGCTTTTACAGATCTTCTGTGAATCTCGTTGGAAGTGGAAATGGAGGTTTATACAAATAA 1587
DB 225 TTGCTTTTACAGATCTTCTGTGAATCTCGTTGGAAGTGGAAATGGAGGTTTATACAAATAA 166

QY 1588 GATACCTCAGTTTTGAATGGTTTTAA 1613
DB 165 GATACCTCAGTTTTGAATGGTTTTATA 140

RESULT 2
BE058404
LOCUS
DEFINITION

BE058404 612 bp mRNA linear EST 14-JUL-2004
sn1507.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl016-11582 5' similar to TR:Q92P22 Q92P22 APS REDUCTASE
PRECURSOR. i, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE058404.1 GI:8402770
EST.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 612)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE
JOURNAL
COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES
source

Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1166 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 433.
Location/Qualifiers
1. 612
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-11582"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl016"
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dr) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelding."

ORIGIN

Query Match 36.8%; Score 599.6; DB 2; Length 612;
Best Local Similarity 99.2%; Pred. No. 9.9e-155;
Matches 602; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 782 AACCTGTGCAATGTGAAGGCCATGACATATGGACTTCCTTAGGACCATGATGTG 841
DB 4 AGCCCTGTGCAATGTGAAGGCCATGACATATGGAACTTCCTTAGGACCATGATGTG 63

QY 842 CCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCATAGG 901

```

Db      64  CCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACATAGG 123
Qy      902  CCTGTTTACCTGGCAACATGAAAGGGAAGGAGGTGGTGGTGGGAGGATGCCAAAGCT 961
Db      124  CCTGTTTACCTGGCAACATGAAAGGGAAGGAGGTGGTGGTGGGAGGATGCCAAAGCT 183
Qy      962  AAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGA 1021
Db      184  AAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGA 243
Qy      1022  AATGGGCTATCCCAATCCCAATGCAATGCTGATGCTACCACTGCTGCTGATGCAATTTCAAC 1081
Db      244  AATGGGCTATCCCAATCCCAATGCAATGCTGATGCTACCACTGCTGCTGATGCAATTTCAAC 303
Qy      1082  AGCCCGAATGTAGTTAACTTTGAGCAGGACTGGAAATGAGAAATTTGGCAAAATTTGGAGGAC 1141
Db      304  AGCCCGAATGTAGTTAACTTTGAGCAGGACTGGAAATGAGAAATTTGGCAAAATTTGGAGGAC 363
Qy      1142  CGAAAGGAACATGGCTTGTGCTTTATGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201
Db      364  CGAAAGGAACATGGCTTGTGCTTTATGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
Qy      1202  GAGGAATCTTATGTTAGCTTTAGCAGACAAGTTAGCAGGCTCAACAGGATGAAAGTTGGA 1261
Db      424  GAGGAATCTTATGTTAGCTTTAGCAGACAAGTTAGCAGGCTCAACAGGATGAAAGTTGGA 483
Qy      1262  AATTTAGAGCAGATGAGAAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGC 1321
Db      484  AATTTAGAGCAGATGAGAAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGC 543
Qy      1322  TTCCTACGATATATTTTCCCAAGCATGCTGCTCGGCCAACAAATTAAGTATCCCTCA 1381
Db      544  TTCCTACGATATATTTTCCCAAGCATGCTGCTCGGCCAACAAATTAAGTATCCCTCA 603
Qy      1382  GAAAGA 1388
Db      604  GAAAGA 610

RESULT 3
CNSADUO
LOCUS   Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION   GSLTFB472C07 of Flowers and buds of strain col-0 of Arabidopsis
            thaliana (thale cress).
ACCESSION   BX813913
VERSION     BX813913.1 GI:42474399
KEYWORDS    HTC; GSUT_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1 (bases 1 to 1535)
            Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
            Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1535)
            Genoscope.
AUTHORS     Direct Submission
TITLE       Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

```

```

Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
source
1..1535
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB472C07"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1535)
/gene="At1g62180"

```

ORIGIN

```

Query Match      35.0%; Score 570.6; DB 4; Length 1535;
Best Local Similarity 68.5%; Pred. No. 1.4e-146;
Matches 882; Conservative 0; Mismatches 359; Indels 46; Gaps 5;

Qy 157 CAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACGCCGA---ACCTCCACGCAA 213
Db 146 CCATTTGTCTCAGAGAGGTTTACTCTATGAAACCTTTAAACGCTGAGTCACATTCCACGAAG 205
Qy 214 GGATTCATTTGCTCTCGCAGCAACCAACATCGTTGCTTCTGCTTCTGAGA-----C 267
Db 206 CGAATCTTGGGTACTCGTCTTCAACTCTAAATGCTCTGAAAGTTGAAGAGGAAGGAGG 265
Qy 268 GAAAGAGGAGAGATTTTGAACAGATAGCCAGTATCTCGAATGCTTCACTCTTGAAT 327
Db 266 AGAAGTTGAAGACTTTGAGCAACTTGTCTAAAGCTTGAAGATGCTTCTCACTTGAAT 325
Qy 328 CATGGATAGAGCCCTCGACAAATTCGGCAACCGATAGCTATTGCTTCACTGAGTGTCTGA 387
Db 326 CATGGATTAAGCTCTTGGAGAGATTCGGAGACCAAAATCGCAATTTGCTTTTAGTGAGCTGA 385
Qy 388 AGATGTTGCTTGAATGAGTATCGGAAATTTGAGGGTCTGACCCCTTTAGGGTTTTCAGTTT 447
Db 386 AGATGTTGCAATGATTAATGATATGACGTTTAACTGGAAAGCCATTTAGGGTTTTCAGTTT 445
Qy 448 GGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTTCAGAGCAATTA 507
Db 446 AGATCAGGAGAGATTAACCCCTGAAACGCTACAGGCTCTTTGACGCGAGTCGAGAGCAAGTA 505
Qy 508 TGGAAATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTGGTGAGGAG 567
Db 506 CGGGATTTCGAATTTGAGTACATGTTTCTGATGAGTTGAGTTTCAAGCTTTTAGTGAGGAA 565
Qy 568 TAAGGGGTTATTTCTTCTTCTACAGAGATGGGCAACAGAGTGTTCAGAGGTGTGAGAAAGGT 627
Db 566 CRAAGGGTTTGTCTTCATTTCTATGAAGATGCTCATCAAGAGTGTTCGCGTGTGAGGAAAGT 625
Qy 628 GAGGCCCTTTAAGGAGGCGCTTAAGGCTCTCAGAGCATGGAATACTGCTCAGAGGAAGA 687
Db 626 TAGACCTTTGCGTGTGCTCTTAAGGGTCTTTAAGGCTTGAATTTACAGGACAGAGGAAGA 685
Qy 688 CCAGTCACCTGTGTAGGTCTGAAATAACCGGTTGTTTTCAGGTTTGATCCCGCTTTTGAGGG 747
Db 686 CCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCAGTTGATCCAGTGTGTTGAAGG 745
Qy 748 AATGGATGTGGAATTTGGAAGCTTGGTGAAGTGGAACCCCTGTTGCAAAATGTGAGAGGCCA 807
Db 746 GTTAGATGGCGGTGTGGAAGTCTTGTGAAGTGAATTCCTTTGCGCTAAATTTGAAGAGGC 805
Qy 808 TGACATATGGAATCTCTTATGAGCAACATGAATGTCCTGTGAAATTCCTTGCATGCAAAAGG 867
Db 806 TGATGTGTGAACCTTTCTGAGAACTATGAGATTTCCGGTGAATGCAATTCGCGCAAGG 865

```

```
Qy 868 ATATGTTTCCATTTGGGTGTGAGCCCTGCATAGCCCTGTTTACCTGGGCAACATGAAG 927
Db 866 GTATGTGTCAATCGGTTGTGAGCCGTGTACTAGCCCGTCTTCCAGGCCCAACATGAGAG 925
Qy 928 GGAAGGAGGTGGTGGTGGGAGGATGCAAAAGCTTAAGCAATGTGCTTCCACAAAGGAAA 987
Db 926 AGAAGGAAGTGGTGGTGGGAGATGCTAAAGCTTAAGATGTGTTCTACACAAAGGGAA 985
Qy 988 TGTAAACACGACGAGAAAGAGGAGATGTTAATGGAAATGGGCTATCCCAATCCCATGCAAA 1047
Db 986 CATCAAGGAGGAGATGGTG-----CTGCAGA 1012
Qy 1048 TGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTAACTTGAAGCAG 1107
Db 1013 CTCAAGCCCTGCTGCTGTGCAAGAGATATTGGAAGCAACAAATGTGTGTCATTGAGCAA 1072
Qy 1108 GACTGGAAATTTGAGAAATTTGGCAAAATTTGGAGGACCCGAAAGCAATGATGGCTTGTGTGCT 1167
Db 1073 AGNAGGGGTGTGAGAACTTTTGAAGCTAGAGAACCGTAAAGAGCGGTGTGTGCTGCTACT 1132
Qy 1168 TTATGCACCATGTGTCCTCTACCTCCAGGCTATGAGGAAATCTTATGTTGACTTTAGCAGA 1227
Db 1133 TTACGCTCTTGTGTGCTCTTCTGCCAGGCGATGGAAGCATCGTACATCGAAATTTGGCTGA 1192
Qy 1228 CAAGTTAGCAGGTCACACGGATGAAGTTTGGAAATTTAGACAGATCGAGAACAGAA 1287
Db 1193 GAAACTTTCGCGG---AAAGAGAGTTAAAGTGGCGAAATTCGAGCTGACGCTGAGCAGAA 1249
Qy 1288 AGAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATA-TTATTTTTCCTCAA 1346
Db 1250 GGAGTTTGTCTAAGCAAGCTTCAGTTAGGAGCTTCCCAAGATATCTGCTCTTTCGCG 1309
Qy 1347 AGCATTCGTTCGCGCCCAACATTAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGA 1406
Db 1310 AAAAGAGCTGCGCAGGCTATTAAAGTACCCTTCAGAGCATAGATATGCTGATTCACCTCA 1369
Qy 1407 TGGCATTTGTAATGCGCTTAAGATGAG 1433
Db 1370 TGTGCTTTGTAATCGTCTTCGCTGAG 1396

RESULT 4
CNSOACVY 1579 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTSL682B02 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX815501
BX815501.1 GI:42473038
HTC; GSLT_cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1579)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1579)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
```

```
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1579
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL682B02"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1579)
/gene="Atlg62180"

gene
ORIGIN
Query Match 34.8%; Score 567.2; DB 4; Length 1579;
Best Local Similarity 68.4%; Pred. No. 1.3e-145;
Matches 880; Conservative 0; Mismatches 358; Indels 48; Gaps 5;

Qy 157 CAATTATCTCAAGACGGAGCTTGGTAAAGCCGTTAAACGCCGA---ACCTCCACCAA 213
Db 209 CCATTGTCTCAGACAGCTTACTCTATGAGACCTTTAAACGCTGAGTCACATTCACGAAG 268
Qy 214 GGATTCATTTGTTCTCTCCGACCAACAACATCGTTGCTTCTGCTTCTGAGA-----C 267
Db 269 CGAATCTGGGTACTCTGCTCTCACTCTAATGTCTCTGAGTTGAAGTGAAGTGAAGTGAAGT 328
Qy 268 GAAAGAGGAAGATTTTGAACAGATAGCCAGTGTCTCGACAAATGCTTCACTCTTGAAGTGAAG 327
Db 329 AGAAGTTGAAGACTTTGAGCAACTTGTCTAAGGAGCTTGAAGATGCTTCTCCACTTGAGAT 388
Qy 328 CATGATAGAGCCCTCGACAAATTCGCAACAGCATAGCTATTCGCTTCAGTGTGCTGA 387
Db 389 CATGATAGAGCTCTTGAGAGATTCGGAGACCAATCGCAATGCTTCTTGTAGTGAGCTGA 448
Qy 388 AGATGTTGCTTTGATTGAGTATGCAAAATTCAGCGGTTCGACCCCTTTAGGGGTTTTCAGTTT 447
Db 449 AGATGTTGCTTTGATTGAGTATGCAAAATTCAGCGGTTCGACCCCTTTAGGGGTTTTCAGTTT 508
Qy 448 GGACACTGGGAGACTGAACCCAGAAACTTATCAAACTTTTGTATGCGGTTCGAGAGCATTA 507
Db 509 AGATACAGGGAGATTAACCCCTCGAGACGTACAGGCTCTTTGACGCGAGTCGAGAAAGCAGTA 568
Qy 508 TGGAAATTCGATTCAGTACATGTTCCCTGATGCTGTGAGGTTCCAGGATTCAGGATTCGAGGAG 567
Db 569 CGGATTCGAAATTCAGTACATGTTTCTGATGCGAGTTGAGGTTCAAGCTTTAGTGAGGAA 628
Qy 568 TAAGGGGTATTCTCTTTTACGAGGATGGCGACCAAGAGTGTTCGAGGGTTCAGAAAGGT 627
Db 629 CAAAGGTTGTTCTCTCTCTCTATGAGATGTGTCATCAAGAGTGTGCCGTGTGAGGAAAGT 688
Qy 628 GAGGCTTTAAGGAGGCGCTTAAAGGTTCTCAGAGCATGGAATACTGCTCAGAGGAAAGA 687
Db 689 TAGACCTTTGCGTCTGCTCTTTGAGGGTCTTAAAGCTTGGATTGAGTACAGGACAGAGGAAAGA 748
Qy 688 CCAGTCACTGGTACTAGGTTGAAATACCGGTTGTTTCAGGTTGATCCGGCTTTTGAGGG 747
Db 749 CCAATCTCCGGGTACGAGATCTGAGATCCCTTATTTGTTTCAGGTTGATCCAGTGTTTGAGAGG 808
Qy 748 AATGCGTGTGGAATTTGGAAGCTTTGGTCAAGTGGAAACCTCTGTCGAAATGTGAAGGGCCA 807
Db 809 GTTAGATGGCGGTGTGGAAGTCTTGTGAAGTGGAAATCTTTGGCTAATGTTGAAGGAGC 868
Qy 808 TGACATATGGAATTCCTCTAGGACCATGAATGTGCTGTGAATTCCTTGCATGCAAAAGG 867
Db 869 TGATGTGTGGAACCTTCTCTGAGAACTACGAGATGTTCCGGTGAATGCAATGTCACGCAAGG 928
```

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

1..1579
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL682B02"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1579)
/gene="Atlg62180"

gene

ORIGIN

Query Match 34.8%; Score 567.2; DB 4; Length 1579;
Best Local Similarity 68.4%; Pred. No. 1.3e-145;
Matches 880; Conservative 0; Mismatches 358; Indels 48; Gaps 5;

QY 868 ATATGTTTCCATTGGGTGAGCCCTGCACTAGGCTGTTTACCTGGCAACATGAAG 927
 Db |||||
 QY 929 GTATGTGTCAATCGGTGTGAGCGGTGTACTAGCGCGTGTCTCCAGSCCAACATGAG 988
 Db |||||
 QY 928 GGAAGGAGGTGTGTGGGAGGATGCAAGCTTAAGATATGGTCTTCAAAAGAAA 987
 Db |||||
 QY 989 AGAAGGAAGTGTGTGGGAGAGTCTAAAGCTTAAGATATGGTCTTACAAAGGAA 1048
 Db |||||
 QY 988 TGTAAAGCAGCAGAAAGAGAGAGTGTAAATGGAATGGGTATCCCAATCCCATGCAA 1047
 Db |||||
 QY 1049 CATCAAGAGGAGATGGTG-----CTGCAGA 1075
 Db |||||
 QY 1048 TGTGATGTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAG 1107
 Db |||||
 QY 1076 CTCAAAGCCTGTGCTGTGCAAGAGATATTTGAAAGCAACAATGTGGTTGCAATTGAGCAA 1135
 Db |||||
 QY 1108 GACTGGAATTTGAGAAATTTGCAAAATTTGAGAGCCGGAAGAACCATGGCTTGTGCT 1167
 Db |||||
 QY 1136 AGAGGGGTGTGAGAACTTTTGAAGCTTAGAGAACCGTAAAGAGCGGTGTGCTGCTACT 1195
 Db |||||
 QY 1168 TTATGACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
 Db |||||
 QY 1196 TTACGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
 Db |||||
 QY 1228 CAAGTTAGCAGGTTCAACAGGATGAAGTTGGAATTTTAGAGCAGATGGAGAACAGAA 1287
 Db |||||
 QY 1256 GAAACTTGGCGG---AAAAGGAGTTAAAGTGGCGAAATTCGAGCTGACGCTGAGCAGAA 1312
 Db |||||
 QY 1288 AGAATTTGCAAGAGTGAACCTGCAATTTGGAAGCTTCCCTACGATATATTTTCCAAA 1347
 Db |||||
 QY 1313 GGAGTTTGTGTAAGCAAGAGCTTCAGTTAGGAGCTTCCGACGATATCTTATTTCCGAA 1372
 Db |||||
 QY 1348 GCATTGCTCTCGGCCAACATAAGATATCCCTCAGAAAGAGAGATGTTGATTCCTTGAT 1407
 Db |||||
 QY 1373 AAGAGCTCCAGG---GCTATTAGTACCTTTAGAGCATAGAGATGTTGATTCATCAT 1429
 Db |||||
 QY 1408 GGCAATTTGTAATGCCTTAAAGATGAG 1433
 Db |||||
 QY 1430 GTCGTTTGTGAATCTTCTTCGGTGAG 1455
 Db |||||

RESULT 5

BF070490

LOCUS

DEFINITION

BF070490 608 bp mRNA linear EST 13-JUL-2004
 st20h11.v1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl065-1918 5' similar to TR:Q9ZP22 Q9ZP22 APS REDUCTASE
 PRECURSOR. ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 608)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, V., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Insert length: 316 Std Error: 0.00 High quality sequence stop: 386.

FEATURES

source

1..608
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl065-1918"
 /issue_type="germinating shoots"
 /lab_host="DH10B"
 /clone_lib="Gm-cl065"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 33.3%; Score 542.2; DB 2; Length 608;
 Best Local Similarity 94.3%; Pred. No. 8.5e-139;
 Matches 562; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 8 GAGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTCACTTTCAATTTCCGACCA 67
 Db |||||
 QY 12 GAGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTCACTTTCAATTTCCGACCA 71
 Db |||||
 QY 68 ACTTCCACCTTCCCATCATCGNAACCCAACTTCCGAAATTTGGTCAATTAGATTTCG 127
 Db |||||
 QY 72 ACTTCCACCTTCCCATCATCGNAACCCAACTTCCGAAATTTGGTCAATTAGATTTCG 131
 Db |||||
 QY 128 GAGAGCCCATTTGGAGCGCGGTTAATTTCAATTTCTCAAAAGACGAGCTTGGTAAAG 187
 Db |||||
 QY 132 GAGAGCCCATTTGGAGCGCGGTTAATTTCAATTTCTCAAAAGACGAGCTTGGTAAAG 191
 Db |||||
 QY 188 CCGTTTAAACCGCAACCTCCAGCAAGGATTCATTTGTTCTCTCGCAGCAACCAATC 247
 Db |||||
 QY 192 CCGTTTAAACCGCAACCTCCAGCAAGGATTCATTTGTTCTCTCGCAGCAACCAATC 251
 Db |||||
 QY 248 GTTGCTTCTGCTTCTGAGAGGAGGAGAGATTTTGAACAGATAGCCAGTGTCTCGAC 307
 Db |||||
 QY 252 GTTGCTTCTGCTTCTGAGAGGAGGAGGATTTTGAACAGATAGCCAGTGTCTCGAC 311
 Db |||||
 QY 308 AATGCTTCACTCTTCAAAATCATGATAGAGCCCTCGCAAAATTCGCAACGACATAGCT 367
 Db |||||
 QY 312 AATGCTTCACTCTTGAATCATGATAGAGCCCTCGCAAAATTCGCAACGACATAGCT 371
 Db |||||
 QY 368 ATTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
 Db |||||
 QY 372 ATTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 431
 Db |||||
 QY 428 CCGTTTAGGTTTTCAGTTTGGACATCTGGAGCTGAAACCCGAAACTTATCAACTTTT 487
 Db |||||
 QY 432 CCGTTTAGGTTTTCAGTTTGGACATCTGGAGCTGAAACCCGAAACTTATCAACTTTT 491
 Db |||||
 QY 488 GATGCGGTTGAGAAGCATTATGAAATTCGATTGAGTACATGTTCCCTGATGCTGTTGAG 547
 Db |||||
 QY 492 GATGCGGCTGAGAAGCATTATGAAATTCGATTGAGTACATGTTCCCTGATGCTGTTGTC 551
 Db |||||
 QY 548 GTTACGGCATTTGTTGAGGAGTAAAGGGTTATTCTCTTCTACGAGGATGGGCACCA 603
 Db |||||

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Insert length: 316 Std Error: 0.00 High quality sequence stop: 386.

Dd		1	GCTTCACTCTTGAAATCATGGATAGACGCGCTCGACAATAATTCCGCACACGACATAGCTATT	60
Qy		371	GCCTTCAGTGGTGCTGAAGATGTGTCTTTGATTTAGTAGTATGCGAAAATTGACGGFTGCACC	430
Dd		61	GCCTTCAGTGGTGCTGAAGATGTGTCTTTGATTTAGTAGTATGCGAAAATTGACGGFTGCACC	120
Qy		431	TTTTAGGGTTTTCAGTTTTGACACTGGGAGACTGAAACCCAGAAAACTTATCAACTTTTGTAT	490
Dd		121	TTTTAGGGTTTTCAGTTTTGACACTGGGAGACTGAAACCCAGAAAACTTATCAACTTTTGTAT	180
Qy		491	GCSTTTTGAGAAGCATTATGGAATTTCCGATTTAGTAGTACATGTTTCCCTGATGCTGTGAGGTT	550
Dd		181	GCSTTTTGAGAAGCATTATGGAATTTCCGATTTAGTAGTACATGTTTCCCTGATGCTGTGAGGTT	240
Qy		551	CAGGCATTGCTGAGGAGTAAGGGGTTTATCTCTTTTCTACGAGATGGCACCAAGAGTGT	610
Dd		241	CAGGCATTGCTGAGGAGTAAGGGGTTTATCTCTTTTCTACGAGATGGCACCAAGAGTGT	300
Qy		611	TGCAGGGTGAGAAAAGGTGAGGCCCTTTAAGAGAGGGCCCTTAAGGGTCTCAGAGCATGGATA	670
Dd		301	TGCAGGGTGAGAAAAGGTGAGGCCCTTTAAGAGAGGGCCCTTAAGGGTCTCAGAGCATGGATA	360
Qy		671	ACTGGTTCAGAGGAAAGACACGACTCAGCTGGTACTAGGTCTGAAATACCGGTTGTTTCAGGTT	730
Dd		361	ACTGGTTCAGAGGAAAGACACGACTCAGCTGGTACTAGGTCTGAAATACCGGTTGTTTCAGGTT	420
Qy		731	GATCCGGCTTTTTCAGGGAAATGGATGGTGGAAATTTGGAAGCTTGGTGAAGTGGAACCCCTGTT	790
Dd		421	GATCCGGCTTTTTCAGGGAAATGGATGGTGGAAATTTGGAAGCTTGGTGAAGTGGAACCCCTTAT	480
Qy		791	GCAAAATGTGAAGGGCCATGACATATGAGAACTTCTCTTAGGACCATGAATGTGCCTGTGAAT	850
Dd		481	GCAAAATGTGAAGGGCCATGACATATGAGAACTTCTCTTAGGACCATGAATGTGCCTGTGAAT	540
Qy		851	TCCTTGCATGC 861	
Dd		541	TCCTTGCATGC 551	
RESULT 7				
B1316361				
LOCUS				
DEFINITION				
3'-PHOSPHODENOSINE 5'-PHOSPHOSULFATE REDUCTASE ; mRNA sequence.				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R. Public Soybean EST Project Unpublished (1999)				
TITLE				
JOURNAL				
COMMENT				
Contact : Shoemaker R/Public Soybean EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave Brookings, SD				

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 396.

FEATURES

Location/Qualifiers
1..516
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1065-4337"
/tissue_type="germinating shoots"
/lab_host="DH10B"

/clone_lib="Gm-c1065"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 31.3%; Score 510.2; DB 3; Length 516;
Best Local Similarity 99.4%; Pred. No. 6.2e-130;
Matches 512; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
QY 21 ACAGCTAGTTAATGAGCCCTCGCTTTCACCTTCTCAATTTCCGACCAACCTTCCACCTTC 80
DB 2 ACAGCTAGTTAATGAGCCCTCGCTTTCACCTTCTCAATTTCCGACCAACCTTCCACCTTC 61
QY 81 CATCATCGGAACCAACTTCCGCAATTTGGGTCAATTAGGATTTGGAGAGGCCCATTTG 140
DB 62 CATCATCGGAACCAACTTCCGCAATTTGGGTCAATTAGGATTTGGAGAGGCCCATTTG 121
QY 141 GAGCGCGGTTAATTTCAATTTATCTCAAGAGCGAGCTGGTAAAGCCGTTAAAGCCCG 200
DB 122 GAGCGCGGTTAATTTCAATTTATCTCAAGAGCGAGCTGGTAAAGCCGTTAAAGCCCG 181
QY 201 AACCTCCAGCAGGATTCATTTGCTCTCCGCAACCAACCATCGTTGCTTCTGCTT 260
DB 182 AACCTCCAGCAGGATTCATTTGCTCTCCGCAACCAACCATCGTTGCTTCTGCTT 241
QY 261 CTGAGACGAAAGAGGAAGATTTTGAACAGATAGCCAGTGATCTCGACAATGCTTCACTC 320
DB 242 CTGAGACGAAAGATGAAGATTTTGAACAGATAGCCAGTGATCTCGACAATGCTTCACTC 301
QY 321 TTGAATATCATGTAGAGCCCTCGACAAATTTGGCAACGACATAGCTATTTGCTTCAGTG 380
DB 302 TTGAATATCATGTAGAGCCCTCGACAAATTTGGCAACGACATATCTATTTGCTTCAGTG 361
QY 381 GTCTCTGAAGATTTGCTTTGATTTAGTATGCGAAATTTGACGGGTGACCCCTTTAGGGTTT 440
DB 362 GTCTCTGAAGATTTGCTTTGATTTAGTATGCGAAATTTGACGGGTGACCCCTTTAGGGTTT 421
QY 441 TCAGTTTGACACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTTGAGA 500
DB 422 TCAGTTTGACACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTTGAGA 481
QY 501 AGCATTTAGGAATTCGATTTAGTATGATGTTTCCT 535
DB 482 AGCATTTAGGAATTCGATTTAGTATGATGTTTCCT 516
```

RESULT 8
CB339330
LOCUS

CB339330 830 bp mRNA linear EST 14-MAR-2003

DEFINITION

ACCESSION CB339330
VERSION CB339330.1
KEYWORDS GI:28958671
SOURCE Vitis vinifera
ORGANISM Vitis vinifera

REFERENCE

AUTHORS Jones, K., Walker, M.A. and Cook, D.R.

TITLE

Transcriptional responses of Vitis vinifera to infection by the bacterial pathogen Xylella fastidiosa

JOURNAL

Unpublished (2003)

COMMENT

Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617

Email: drcocook@ucdavis.edu

Seq primer: ACGGTACCGGACATATGCC.

FEATURES

source

1..830
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA23EI031VPB_B07"
/sex="Hexamphrodite"
/dev_stage="Mid-season leaf material, collected July 25, 2001."
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Leaf - CA23EI"
/note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2: SfiI; CA23EI is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were verified to be infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AACCTAGGTCATCAACGAGTGGCCATTACGCCGGG-3' and
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 30.9%; Score 504; DB 6; Length 830;
Best Local Similarity 76.7%; Pred. No. 3.7e-128;
Matches 647; Conservative 0; Mismatches 179; Indels 18; Gaps 2;

```
QY 270 AAGAGGAAGATTTTGAACAGATAGCCAGTATCTCGACATGCTTCACCTCTTGAATCA 329
DB 4 AGGTGGAGGACTATGAGCAATTTGGCATGACATTTGCCATTTGCTTCAATTA 63
QY 330 TGGATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGCTTTCAGTGGTCTGAAG 389
DB 64 TGGATAGAGCCCTCGCAAAATTTGGCAATGACATTTGCCATTTGCTTCAATTA 123
QY 390 ATGTTGCTTTGATGATGCGAAATTTGACGGGTGACCCCTTTAGGGTTTTCAGTTGG 449
DB 124 ATATTGCTTTGATGATGCGCTTAACCTGCTCGTTCGTTTATGCGGTATTCAGCCTGG 183
QY 450 ACACCTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTGAGAACGATATG 509
DB 184 ACACCTGGAGGTTGAACCCGGAACATATCAATTTCTTTGACCGTTGAGAAACACATATG 243
```

```
Qy 510 GAATTCGCATTGATCATGTTCCCTGATGCTGTGTGAGTTTCAGGCATTGGTGAGAGTA 569
Db 244 GCATCCGATTTGATATCATGTTTCCAGATGCTGTGTGAAGTTTCAGGCTTGGTGAGGACA 303
Qy 570 AGGGGTTTATTTCTCTTCTACGAGGATGGCCACCAAGAGTGTTCAGGCTGAGAAAGTGA 629
Db 304 AGGAGCTGTTTCTCTCTACGAGGATGGCCACCAAGAGTGTTCAGGCTGAGAAAGTGA 363
Qy 630 GGCCTTTAAGAGGGCCCTTAAGGCTCTCAGAGCATGATAACTGCTCAGAGGAAAGACC 689
Db 364 GACCTTTGAGAGGGCCACTGAAGGTTTGGCTGCTGGATCACAGGGCAAGAGAGATC 423
Qy 690 AGTCACCTGTTACTGAGTCTGAATACACCGTGTGTTCAGGTTGATCCGGCTTTTGAGGCAA 749
Db 424 AGTCTCTGTGTACAGGGCTGAAGTCTGTGTTCAGGTTGATCCAGGCTTCAGGGGT 483
Qy 750 TGGATGTTGGAATTTGGAAGCTTGGTGAAGTGGAAACCCCTGTTGCAAAATGTGAAGGGCCATG 809
Db 484 TGGATGTTGAGGCTTGGCAGCCTGGTGAATGGAACCCAGTGGCAAAATGTGCAGGGCATGG 543
Qy 810 ACATATGGAATCTCTTAGACCATGATGTCCTGTAATCTTCATGTCGCAAAAGAT 869
Db 544 ACATCTGGAATTTCTCCGCTGATGAATGTCCTGTAATCTTCATGTCGCAAAAGAT 603
Qy 870 ATGTTTCCATTTGGGTGTGAGCCTGCACTAGGCTGTTTACCTGGGCAACATGAAAGGG 929
Db 604 ATATTTCAATTTGGGTGTGAGCCTGCACTAGGCTGTTTACCTGGGCAACATGAGAGAG 663
Qy 930 AAGGAGGTGTGTGGGAGGATCCCAAGCTAAGGAATGTGTCCTTACCAAAAGGAATG 989
Db 664 AAGGAAGATGTTGTTGGGAGATCCCAAGGCTAAGGAATGTGTCCTTACCAAAAGGAATG 723
Qy 990 TAAAGCAGCAAGAGAGGATGTTAATGGAATGGCTATCCCATCCCATCAATG 1049
Db 724 TCA---AGCAGGAAGATGGAACAAAATGGAATGGGCATGCAAA----- 766
Qy 1050 GTGATGTACACATGTGCTGACATTTTCAACAGCCGGAATGTAGTTAACTTGACGAGGA 1109
Db 767 -TGAACCGCCACATGTTCTGATCTCTTCGACACCCAGNNNTGTTACCTTAACGAGGA 825
Qy 1110 CTGG 1113
Db 826 CTGG 829
```

```
RESULT 9
B1316488
LOCUS
DEFINITION
B1316488 507 bp mRNA linear EST 21-JUL-2004
ID: Gm-cl065-4568 5', similar to TR:Q39619 Q39619
PAPS-REDUCTASE-LIKE PROTEIN. ; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 507)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Waller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
```

Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 410.
Location/Qualifiers
1. 507
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-4568"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 30.2%; Score 491.6; DB 3; Length 507;
Best Local Similarity 98.2%; Pred. No. 9e-125;
Matches 497; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 21 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTTCCTTCGACCACTTCCACCTTCC 80
Db 2 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTTCCTTCGACCACTTCCACCTTCC 61
Qy 81 CATCATCGAACCCAAACTTCCGCAATTTGGTCAATTTAGGATTTCCGAGAGGCCATTG 140
Db 62 CATCATCGAACCCAAACTTCCGCAATTTGGTCAATTTAGGATTTCCGAGAGGCCATTG 121
Qy 141 GAGCGCCGTTTAAATTTTCAATTTATCTCAAGACGAGCTTGGTAAAGCCGTTAACGCCG 200
Db 122 GATTCCTTTTAAATTTTCAATTTATCTCAAGACGAGCTTGGTAAAGCCGTTAACGCCG 181
Qy 201 AACCTCCAGCAGGATTCATTTGTTCTCTCGCAGCAACACCATCGTTCTTCTGCTT 260
Db 182 AACCTCCAGCAGGATTCATTTGTTCTCTCGCAGCAACACCATCGTTCTTCTGCTT 241
Qy 261 CTGAGACGAAGAGGAAGATTTTCAACAGATAGCCAGTGTCTCGACAAATGCTTCACCTC 320
Db 242 CTGAGACGAAGAGGAAGATTTTGAACAGATAGCCAGTGTCTCGACAAATGCTTCACCTC 301
Qy 321 TTGAAATCATGGATAGAGCCCTCGACAAATTTGGGAAACGACATAGCTATTGCTTTCAGTG 380
Db 302 TTGAAATCATGGATAGAGCCCTCGACAAATTTGGGAAACGACATAGCTATTGCTTTCAGTG 361
Qy 381 GTGCTGAAGATGTTGCTTTCGATGATGCGAATTTGACGGGTGCGCCCTTTAGGGTTT 440
Db 362 GTGCTGAAGATGTTGCTTTCGATGATGCGAATTTGACGGGTGCGCCCTTTAGGGTTT 421
Qy 441 TCAGTTTGGACACTGGGAGACTGAACCCAGAAAATCTTATCAACTTTTGTATCGGTTGAGA 500
Db 422 TCAGTTTGGACACTGGGAGACTGAACCCAGAAAATCTTATCAACTTTTGTATCGGCTGAGA 481
Qy 501 AGCATTTATGGAATTTGCAATTTGAGTAC 526
|||||

Db	482	AGCATTATGGAAATTCGCATTGAGTAC	507	
<p>CK264757 959 bp mRNA linear EST 03-AUG-2004</p> <p>EST710835 potato abiotic stress cDNA library Solanum tuberosum cDNA</p> <p>clone POABU61 5' end, mRNA sequence.</p> <p>CK264757</p> <p>CK264757.1 GI:39821735</p> <p>EST.</p> <p>Solanum tuberosum (potato)</p> <p>Solanum tuberosum</p> <p>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum.</p> <p>1 (bases 1 to 959)</p> <p>Buell, C. R., Hart, A., Zismann, V., Karamycheva, S. A. and Baker, B.</p> <p>Generation of ESTs from abiotic stressed potato tissue</p> <p>Unpublished (2003)</p> <p>Other ESTs: EST710836</p> <p>Contact: Robin Buell</p> <p>The Institute for Genomic Research</p> <p>9712 Medical Center Dr, Rockville, MD 20850, USA</p> <p>Email: potato-array@igrr.org</p> <p>Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/.</p> <p>Seq primer: ATT TAG GTG ACA CTA TAG.</p> <p>Location/Qualifiers</p> <p>1..959</p> <p>/organism="Solanum tuberosum"</p> <p>/mol_type="mRNA"</p> <p>/cultivar="Kennebec"</p> <p>/db_xref="taxon:4113"</p> <p>/clone="POABU61"</p> <p>/tissue_type="abiotic stress treated leaf and root tissue"</p> <p>/lab_host="DH10B-Tona"</p> <p>/clone_lib="potato abiotic stress cDNA library"</p> <p>/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."</p>				
<p>ORIGIN</p> <p>Query Match 29.9%; Score 486.6; DB 7; Length 959;</p> <p>Best Local Similarity 73.1%; Pred.No. 2.6e-123;</p> <p>Matches 691; Conservative 0; Mismatches 224; Indels 30; Gaps 4;</p>				
Qy	509	GGAAATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGATTGAGGAGT	568	
Db	1	GGCATTGCGATTGAATACATGTTCCCTGATTCAGTTGAGTTGAGCTTTGTTAGACC	60	
Qy	569	AAGGGGTTATCTCTTTTACGAGATGGGCCAACAGAGTGTTCAGGGTGAGAAAGGTTG	628	

```

Glycine.
1 (bases 1 to 629)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvett,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,I., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Other ESTs: BE823898 corresponding to Gm-r1070-8298 (3')
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1537 Std Error: 0.00
High quality sequence stop: 446.
FEATURES             source
    Location/Qualifiers
        1..629
            /organism="Glycine max"
            /mol_type="mRNA"
            /db_xref="taxon:3847"
            /clones="GENOME SYSTEMS CLONE ID: Gm-r1030-1413"
            /lab_host="DH10B"
            /clone_lib="Gm-r1030"
            /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
            cDNA library was constructed from mRNA isolated from
            immature cotyledons of greenhouse grown plants
            (individual seed fresh weight of 100-300mg). The library
            was prepared using the Life Technologies pSuperScript cDNA
            library construction kit. Complementary DNA was
            synthesized from mRNA using a poly(dT) sequence with a
            NotI restriction site. SalI linkers adapters were ligated
            to the blunt-ended cDNA fragments followed by NotI
            digestion. The cDNA fragments were directionally cloned
            into the NotI-SalI restriction site of the pSPORT1
            vector. The ligated cDNA fragments were transformed into
            E. coli ElectroMax DH10B host cells. This library was
            constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
            that Gm-r1030 is a re-rack of Gm-cl007."
ORIGIN
Query Match          29.6%; Score 481.4; DB 1; Length 629;
Best Local Similarity 85.4%; Pred. No. 6.5e-122;
Matches 536; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 276 AAGATTTTGAACAGATACCCAGTATCTCGAATGCTTCACCTCTTGAATCATGGATA 335
DB 1 AAGATTTTGAACAAATTAGCGAAGACCTTGAAATTCATCTCTTGGAGATTATGGATA 60
QY 336 GAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTCAGTGGTCTCGAAGATGTTG 395
DB 61 AGGCCCTCGAATAATTTGGGAACGACATCGCTATTGCTTTAGTGGTCTCGAAGATGTTG 120
QY 396 CTTTGATTGATGATGCGAAATTTGACGGTTCGACCCCTTTAGGGTTTCAGTTTGGACACTG 455
DB 121 CTTTGATTGATGATGCAATTGTGACGGTTCGACCTTCACAGATGTTTATGCTTGACACTG 180
QY 456 GGAGACTGAACCCAGAAACTTATCAACTTTTGTGATGCGGTTTGAGAAGCAATTATGGAATTC 515
DB 181 GGAGACTGAACCCAGAAACCTTACAAATTTTGTGACGCTGTTGAGAAGCAATTATGGAATTC 240
QY 516 GCATTGACTACATGTTCCCTGATGCTGTGTAGGTTTCAGGCATTTGGTGGAGTAAAGGGT 575
DB 241 ATATTGAGTACATGTTCCCTGATGCGGTTGAGGTTGAGGCATTTAGTAGAACTAAGGGGC 300
QY 576 TATTCTCTTTCTACGAGGATGGGCCAACCAAGAGTGTGTGACGGGTGAGAAAAGGTGAGGCCTT 635

```

```

Db 301 TCTTCTCATTTTACGAGGATGGGCATCAAGAGTGTCTGTAGAGTAAGAAAGGTGAGGCCCT 360
QY 636 TAGGAGGGCCCTTAAGGCTCTCAGAGCATGGATACTGCTCAGAGGAAGAACCCAGTCCAC 695
DB 361 TGAGGAGAGGCCCTTAAGGGGCTCAAAAGCATGGATTACTGGACAGAGAAAAGACCAAGTCTC 420
QY 696 CTGCTACTAGGTCCTGAAATACCCGGTTGTTTCAGGTTGATCCGGCTTTTGGAGGAATGGATG 755
DB 421 CTGGTACTAGGTCCTGAAATCCCTATTGTGCAGGTTGATCTGTTTGGAGGACTGGATG 480
QY 756 GTGGAATTGGAAGCTTGGTGAAGTGGAAACCTCTGTTGCAAAATGTGAAGGGCCATGACATAT 815
DB 481 GTGGAATTGGAAGCTTGGTGAAGTGGAAACCCGGTTGCAAAATGTTAAATGCTGTAGACATAT 540
QY 816 GGAACCTCTTAGGACCATGAATGTGCTGCTGTAATCTCTTGGCATGCAAAAGGATATGTTT 875
DB 541 GGAACCTTACTTAGGACCATGAATGTTCTCTGTAGATTCAATTGCAATCCCAAGGATATGTTT 600
QY 876 CCATTGGGTGTGAGCCCTGCACTAGGCC 903
DB 601 CGATTGGCTGTGAGCCATGATCAAGGC 628

RESULT 12
CB340603/c
LOCUS CB340603 704 bp mRNA linear EST 14-MAR-2003
DEFINITION CA23E1011VR All Cabernet Sauvignon Leaf - CA23E1 Vitis vinifera
cDNA clone CA23E1011VR All 3', mRNA sequence.
ACCESSION CB340603
VERSION CB340603.1 GI:28961185
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
REFERENCE Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Vitaceae; Vitis.
1 (bases 1 to 704)
Goes da Silva,F., Lim H., Iandolino,A., Baek,J., Leslie,A., Xu,J.,
Jones,K., Walker,M.A. and Cook,D.R.
Transcriptional responses of Vitis vinifera to infection by the
bacterial pathogen Xylella fastidiosa
Unpublished (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcocook@ucdavis.edu
Seq primer: GCCAACGAAATGGTCTAG.
FEATURES             source
    Location/Qualifiers
        1..704
            /organism="Vitis vinifera"
            /mol_type="mRNA"
            /cultiivar="Cabernet Sauvignon"
            /db_xref="taxon:29760"
            /clones="CA23E1011VR_All"
            /sex="Hermaphrodite"
            /dev_stage="Mid-season leaf material, collected July 25,
            2001."
            /lab_host="DH5alpha"
            /clone_lib="Cabernet Sauvignon Leaf - CA23E1"
            /note="Organ: Leaf; Vector: pDNR; Site_1: Sfil; Site_2:
            Sfil; CA23E1 is a cDNA library of Cabernet Sauvignon
            leaves. The leaves were collected on July 25, 2001, in
            Napa Valley, California, and represent leaves in
            mid-season development. These leaves were verified to be
            infected with the bacterial pathogen, Xylella fastidiosa,
            based on a diagnostic assay using PCR and Xylella-specific
            primer pairs. The plants were asymptomatic at the time of
            collection, but later developed symptoms. cDNAs were made
            by oligo-dT priming and directionally cloned. 5' and 3'

```


Db 339 GAGCTCTTGAAGATTGGAAATGACATAGCGATTGCTTTCAGTGGTCTGAAGATGTTG 398
Qy 396 CTTTGATGTAGTATGCGAAATATGACGGGTGACCCCTTTAGGTTTTCAGTTTGACACTG 455
Db 399 CTTTGATCGAGTATGCCCAATTAACCTGCTCGACCAATTTAGGCTGTTTGTAGCTGATAC 458
Qy 456 GAGACTGACCCAGCAAACTATCACTTTTTCATGCGGTGAGAGCAATTTATGGAATTC 515
Db 459 GAAGGCTAAACCCAGAAACATATAGATTCCTTTGATGAAGTTGAGAAACATTTATGGCATAC 518
Qy 516 GCATTGATGATCATGTTCCCTGATGCTCTTTCAGGTTTCAGGCAATTTGGTGAGGAGTAAGGGGT 575
Db 519 GTATTGATACATGTTTCTGATGCTGTTGAAGTTTCAGGCAATTTAGTGAGGAGCAAGGGC 578
Qy 576 TATTCTCTTTCTAGGATGGGCAACAGAGTGTTCAGGGGTGAGAAAGTGAGGCTT 635
Db 579 TCTTCTCTTCTATGAGATGGGCACCAAGGATGCTGCCGTGTGAGGAAAGTGAGGCC 638
Qy 636 TAAGGAGGCCCTTAAGGTTCTCAGACATGATTAACCTGCTCAGAGCAAGCAAGCAAGCAAGTAC 695
Db 639 TTGGAGGGCCCTCAAGGATTTGCGTGCCTGGATCACTGGCCAAAGGAAAGATCAGTCTC 698
Qy 696 CTGCTACTAGTCTGAATACCCGTTCTTTCAGGTTGATCCGGCTTTTTCAGGGAATGGATG 755
Db 699 CAGGTACTAGTCTGAATCCCTGATGCTCAGGTTGATCCGTTTTCAGGAGGCTGGAAG 758
Qy 756 GTGGAATTTGAAGTTCGTTGAAGTGAACCTGTTGCAAAATGTGAAGGCCCATGACATAT 815
Db 759 GTGGGTTGGCAGCTGTTGAAGTGAATCCAGTGGCTTAATCGAAGGGAAATGACATTT 818
Qy 816 GGAATTTCTTAGACCATGATGCTGCTGATTCCTTTCAGTGCATGCAAAAGGATATCTT 875
Db 819 GGAATTTCTTCGTAATGATGGGTGCTCCATCAATCTTTCAGTCACTAAAGGGTACATCT 878
Qy 876 CCATTGGGTGAGCCCTGCACTAGGCTGT 906
Db 879 CGATTGGGTGAGCCATGCACAGGCCCGT 909

RESULT 14
CF119175/c
LOCUS
DEFINITION MTU10CS.P14.D12 Aspen stem cDNA Library Populus tremuloides cDNA,
mRNA sequence.
CF119175
VERSION
KEYWORDS
SOURCE
ORGANISM
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 764)
Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
Tsai,C.-J.
Expressed sequence tags from Aspen
Unpublished (2003)
Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.
Location/Qualifiers
1..764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/note="Organ: stem"

ORIGIN
Query Match 27.7%; Score 450.8; DB 6; Length 764;
Best Local Similarity 76.1%; Pred. No. 2.1e-113; Indels 6; Gaps 2;
Matches 583; Conservative 0; Mismatches 177;
Qy 599 CACCAAGAGTGTTCAGGGTGAGAAAGGTGAGGCTTTTAAAGGAGGCGCTTAAAGGGTCTC 658
Db 764 CATCAGAGTGTCTCCGCTGTTTGAAGGTGAGACCCCTTGAAGGGGCTTGAAGGGGCTT 705
Qy 659 AGAGCATGGAATACTGGTTCAGAGGAAGACAGTCACTCTGTACTAGTCTGAAATACCG 718
Db 704 CGGGCTTGGATCACTGGCCAAAGGAAGGATCAATCTCCAGGACGAGGTCTGAAATTTCCA 645
Qy 719 GTTGTTCAGGTTGATCCGCTTTTGAAGGAATGGATGGTGGGCTGGAAGCTTCAAG 778
Db 644 GTTGTTCAGGTTGATCCGCTTTTGAAGGAATGGATGGTGGGCTGGAAGCTTCAAG 585
Qy 779 TGAACCCCTCTTGCATAATGTGAAGGGCCATGACATATGGAACCTTCTTTAGGACCATGAAT 838
Db 584 TGAATCCAATGGCAATGTGAGGACAAAGATCTATGGAAGTTCCTTCGAACCATGGAT 525
Qy 839 GTGCTGTGAATTCCTTTCATGCAAGAGATATGTTTCCATTTGGGTGTGAGCCCTGCACT 898
Db 524 GTGCTGTGAATTCCTTTCATGCAAGAGGATACATCTCTATTTGGTTGCGAGCCTTGACA 465
Qy 899 AGGCTGTTTTACCTGGCAACATGAAAGGGAAGGAGGTGTTGGTGGAGGATGCCAAA 958
Db 464 AGGCCAGTTCCTCGGGCCCAACATGAGAGAGAGGAGGTGGTGGAGGATGCCACA 405
Qy 959 GCTAAGGAATGTGCTTTCACAAAGGAATGTAAAGCAGCAGAAAGAGGAGGTGTTAAT 1018
Db 404 GCTAAGGAATGTGCTTTCATTAAGGAATCTGAAACAGGGTGATGCAAGCCAACTTAAT 345
Qy 1019 GGAATGGGTATCCCAATCCCATGCAAAATGTGATGCTACCACTGTGCTGCACTTTTC 1078
Db 344 GGCAATGGAACCGGGCTGCCCATGCAAAACGG---GGCTGCACTGTGTGCTGATATTTTC 288
Qy 1079 AACAGCCCGAATGTAGTTTAACTTTCAGCAGGACTGGAATTCAGAAATTTGGCAAAATTTGAG 1138
Db 287 AACTCCGGAATTTGGTCAATTTGAGCAGGCCCGGAATCGAACTTATTAATTTAGAG 228
Qy 1139 GACCGAAGGAACCATGGCTTGTGCTTTTATGCACTGTGCTCCCTACTGCCAGCT 1198
Db 227 AACCGAAGGACCATGGCTTGTGCTCTACGCCCATGTGTGCCAATTTTGTCCAGGT 168
Qy 1199 ATGAGGAATCTTATGTTGATAGCAGACAAGTTAGCAGGTCACAGGATGAAGTT 1258
Db 167 ATGGAAGCATCTATGTTGAATTTGGCCGATAAAGTTAGCAGGG---AGTGGAAATAAAGGTG 111
Qy 1259 GGAATAATTAGCAGATGGAGACAGAAAGAAATTTGCAAGAGTGAACCTCAATTTGGGA 1318
Db 110 GGAATAATTAGCAGATGGAGATGGCGATCAGAAAGAAATTTCAAAACAAAGAAATTAACAGCTCGGA 51
Qy 1319 AGCTTTCCTACGATATTTATTTTCCCAAGCAATTCGTCTCGGCCAA 1364
Db 50 AGTTTCCAAACATCTTCTTCTTCCCTTAACACATCACTCTCGCCCAA 5
RESULT 15
CF119189/c
LOCUS
DEFINITION MTU10CS.P14.F03 Aspen stem cDNA Library Populus tremuloides cDNA,
mRNA sequence.
CF119189
VERSION
KEYWORDS
SOURCE
ORGANISM
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 764)
REFERENCE 1 (bases 1 to 764)

AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai,C-J.
TITLE Expressed sequence tags from Aspen
JOURNAL Unpublished (2003)
COMMENT Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources & Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu
FEATURES Location/Qualifiers
source 1. 764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/note="Organ: stem"

ORIGIN

Query Match 27.7%; Score 450.8; DB 6; Length 764;
Best Local Similarity 76.1%; Pred.No. 2.1e-113;
Matches 583; Conservative 0; Mismatches 177; Indels 6; Gaps 2;
QY 599 CACCAAGAGTGTTCAGAGGTGAGAAAGGTGAGGCGCTTTAAGAGGGGCCCTTAAGGGGTCTC 658
DB 764 CATCAGGAGTGTCCCGTGTAGGAGGTGAGACCCCTTGAGGCGGCTCTGAGGGGCTT 705
QY 659 AGAGCATGGATACTGTGTAGAGAAAGACAGTCACTGTGTACTAGGTCTGAAATACCG 718
DB 704 CGGGCCCTGGATCACTGGCCAAAGGAAGGATCAATCTCCAGGACGAGGTCTGAAATTC 645
QY 719 GTTGTTCAGGTTCATCCGGCTTTTGGAGGAATGGATGGTGAATTCGAAGCTTGGTGAAG 778
DB 644 GTTGTTCAGGTGACCCGGTTTTGGAGGATGGATGGTGGGCTGGAAGCCTGATCAAG 585
QY 779 TGAACCCCTGTTCGAAATGGAAGGGCCATGACATATGGAATTCCTTAGGACCATGAAT 838
DB 584 TGAATCCCAATGCAAAATGTTAGGGACAAAGATGTATGGAAGTTCCTTCGAACCATGGAT 525
QY 839 GTCCCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCCT 898
DB 524 GTCCCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCCT 465
QY 899 AGCCCTGTTTCTACCTGGCAACATGAAAGGAGGAGGTGTGGTGGGAGGATGCCAA 958
DB 464 AGCCCAAGTTCCTCCGGCCCAACATGAGAGAGAGGAGGTGTGGTGGGAGGATGCCCA 405
QY 959 GCTAAGGAATGTGCTTTCACAAAGGAAATGTAAAGCAGCAGAGAGAGGAGGATGTTAAT 1018
DB 404 GCTAAGGAATGTGCTTTCATTAAGGAATCTGAACAGGGTGTGAGCCCACTTAAT 345
QY 1019 GGAATGGGCTATCCCAATGCCATGCAAAATGGTGTGATGTACCACTGTGCTGACATTTTC 1078
DB 344 GGAATGGAAACGGGGGTGCCCATGCAAAACGG---GGCTGCCACTGTGCTGATATTTTC 288
QY 1079 AACAGCCGCAATGTAGTAACTTGGCAGGACTGGAATTTGGAATTTGGCAAAATTTGGAG 1138
DB 287 AACTCCGAGAAATTTGGTCAATTTGAGCAGGCCCGGAATCGAGAACTTTATTAATTTAGAG 228
QY 1139 GACCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGTTGGTGCCTTACTGCCAGGCT 1198
DB 227 AACCGAAAGGAACCATGGCTTGTGTGCTTCTACGCCCCCATGGTGCCTAATTTGCCAAGT 168
QY 1199 ATGGAGGAATCTTATGTTGATTTAGCAGACAGTATAGCAGGGTCAACAGGGATGAAGTTT 1258
DB 167 ATGGAAGCATCTTATGTTGAAATTTGGCCGATAGTGTAGCAGGG---AGTGAATTAAGGTG 111
QY 1259 GGAATTTAGACAGATGAGAACAGAGAAATTTGCAAGAGTGTGAATGCAATTTGGGA 1318
DB 110 GGAATTTAGACAGATGAGAACAGAGAAATTTGCAAGAGTGTGAATGCAATTTGGGA 51

QY 1319 AGCTTCCCTACGATATATATTTTCCCAAGCATTTGCTCGGCCAA 1364
DB 50 AGTTTCCCAACAATCCTTTTCTTCCCTTAACACTCATCTCGGCCAA 5

Search completed: February 20, 2006, 22:40:40
Job time : 5994 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:55:42 ; Search time 314 Seconds
(without alignments)
9221.809 Million cell updates/sec

Title: US-10-731-525-7
Perfect score: 1629
Sequence: 1 gcacggaggagaccacata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	3	US-09-720-318A-7
2	904	55.5	1795	3	US-09-720-318A-5
3	538.2	33.0	1827	3	US-09-720-318A-9
4	435.2	26.7	1210	3	US-09-720-318A-3
5	364.2	22.4	1215	3	US-09-720-318A-1
6	199.6	12.3	783	3	US-09-328-352-1063
7	149.4	9.2	810	3	US-09-252-991A-13944
8	84.2	5.2	660	3	US-09-252-991A-13837
9	81	5.0	723	3	US-09-902-540-7978
10	81	5.0	6063	3	US-09-902-540-807
11	73.4	4.5	269223	3	US-09-596-002-41
12	72.6	4.5	756	3	US-09-540-236-1037
13	59	3.6	7218	2	US-08-232-463-14
14	42	2.6	1141	3	US-09-806-708B-22
15	39.2	2.4	578	3	US-09-385-982-465
16	38.6	2.4	1830121	3	US-09-557-884-1
17	38.6	2.4	1830121	3	US-09-643-990A-1
18	38.6	2.4	1830121	3	US-10-158-865-1
19	38	2.3	474	3	US-09-621-976-18033
20	36.6	2.2	1101	3	US-08-984-919A-13
21	36.6	2.2	1101	3	US-08-984-919A-14
22	36.6	2.2	1107	3	US-08-781-420-13
23	36.6	2.2	1107	3	US-08-781-420-14
24	36.6	2.2	1107	3	US-08-874-102-13

C	25	36.6	2.2	1107	3	US-08-874-102-14	Sequence 14, Appl
	26	36.6	2.2	1107	3	US-09-006-595A-13	Sequence 13, Appl
C	27	36.6	2.2	1107	3	US-09-006-595A-14	Sequence 14, Appl
	28	36.6	2.2	1401	3	US-08-984-919A-32	Sequence 32, Appl
C	29	36.6	2.2	1401	3	US-08-984-919A-34	Sequence 34, Appl
	30	36.6	2.2	1407	3	US-08-874-102-32	Sequence 32, Appl
C	31	36.6	2.2	1407	3	US-08-874-102-34	Sequence 34, Appl
	32	36.6	2.2	1410	3	US-08-984-919A-54	Sequence 54, Appl
C	33	36.6	2.2	1410	3	US-08-984-919A-56	Sequence 56, Appl
	34	36.6	2.2	1416	3	US-08-874-102-54	Sequence 54, Appl
C	35	36.6	2.2	1419	3	US-08-874-102-56	Sequence 56, Appl
	36	36.6	2.2	1466	3	US-08-984-919A-10	Sequence 10, Appl
C	37	36.6	2.2	1466	3	US-08-984-919A-12	Sequence 12, Appl
	38	36.6	2.2	1472	3	US-08-781-420-10	Sequence 10, Appl
C	39	36.6	2.2	1472	3	US-08-781-420-12	Sequence 12, Appl
	40	36.6	2.2	1472	3	US-08-874-102-10	Sequence 10, Appl
C	41	36.6	2.2	1472	3	US-08-874-102-12	Sequence 12, Appl
	42	36.6	2.2	1472	3	US-09-006-595A-10	Sequence 10, Appl
C	43	36.6	2.2	1472	3	US-09-006-595A-12	Sequence 12, Appl
	44	36.6	2.2	1488	3	US-08-984-919A-49	Sequence 49, Appl
C	45	36.6	2.2	1488	3	US-08-984-919A-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-09-720-318A-7
; Sequence 7, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; TYPE: DNA
; LENGTH: 1629
; ORGANISM: Glycine max
; US-09-720-318A-7

Query Match	100.0%;	Score 1629;	DB 3;	Length 1629;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1629;	Conservative	0;	Mismatches	0; Indels
				0; Gaps
0;				
Qy	1	GCACGAGGAGAGAACCCATTAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Db	1	GCACGAGGAGAGAACCCATTAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Qy	61	CGACCAACTTCCACCTTCCATCATCGGACCCAACTTCCGCAATTCGGTCAATTAG	120	
Db	61	CGACCAACTTCCACCTTCCATCATCGGACCCAACTTCCGCAATTCGGTCAATTAG	120	
Qy	121	GATTCGAGAGGCCCATTTGGAGGCCGCTTAATTTCAATTATCTCAAGACGGAGCTT	180	
Db	121	GATTCGAGAGGCCCATTTGGAGGCCGCTTAATTTCAATTATCTCAAGACGGAGCTT	180	
Qy	181	GTTAAAGCCGTTAAACCCGCAACTCCACGAGGATTCATTTGTTCTCTCGCAGAAC	240	
Db	181	GTTAAAGCCGTTAAACCCGCAACTCCACGAGGATTCATTTGTTCTCTCGCAGAAC	240	
Qy	241	AACCATGTTGCTTCTGCTTCTCAGACGAGGAGGAATTTTGAACAGATAGCAGTGA	300	
Db	241	AACCATGTTGCTTCTGCTTCTCAGACGAGGAGGAATTTTGAACAGATAGCAGTGA	300	
Qy	301	TCTCGACAATGCTTCACTCTTGAATCATGTAGAGCCCTCGACAAATTCGCAACGA	360	
Db	301	TCTCGACAATGCTTCACTCTTGAATCATGTAGAGCCCTCGACAAATTCGCAACGA	360	

Db 552 AGCAATTAAGAAATTCATATTAGTACATGTTCCCTGATCGGTTGAGTTTCAGGCATTAG 611
Qy 561 TGAGGAGTAAGGGGTTATCTCTCTTACAGGATGGCCACCAAGAGTGTTCAGGGTGA 620
Db 612 TAAGAACTAAGGGGCTCTCTCTATTTACAGGATGGGATCAAGAGTGTCTAGAGTAA 671
Qy 621 GAAAGGTAGGCGCTTTAAGGAGGCGCTTAAAGGTTCTCAGAGCATGGAATCACTGGTCAGA 680
Db 672 GAAAGGTAGGCGCTTTCAGGAGAGCGCTTAAAGGTTCTCAGAGCATGGAATCACTGGCAGA 731
Qy 681 GGAAGACCACTGACCTGCTAGTCTAGTCTGAATACCGTGTTCAGGTTGATCCGCTT 740
Db 732 GAAAGACCACTGCTCTGCTAGTCTAGTCTGAATACCGTGTTCAGGTTGATCCGCTT 791
Qy 741 TTGAGGGAATCGATGGTGGAAATGGAACTTGGTGAAGTGAACCTCTGTTGCAAAATGTA 800
Db 792 TTGAGGGAATCGATGGTGGAAATGGCAGCTTGGTGAAGTGAACCTCTGTTGCAAAATGTA 851
Qy 801 AGGCCATGACATATGAACTCTCTTAGGACCATGAATGTGCTGTGAATTCCTTTCATG 860
Db 852 ATGCTCTAGACATATGAACTCTCTTAGGACCATGAATGTGCTGTGAATTCCTTTCATG 911
Qy 861 CAAAGGATATGTTTTCATTTGGGTGTGAGCGCTGCACTAGGCTGTGTTTACCTGGGCAAC 920
Db 912 CCAAGGATATGTTTTCATTTGGGTGTGAGCGCTGCACTAGGCTGTGTTTACCTGGGCAAC 971
Qy 921 ATGAAAGGAGGAGGTGGTGGGAGGATGTCACCAAGCTAAGGAATGTGTTCTTCA 980
Db 972 ATGAAAGGAGGAGGTGGTGGGAGGATGTCACCAAGCTAAGGAATGTGTTCTTCA 1031
Qy 981 AAGGAAATGTAAGCAGCAAGAGGAGGATGTTAATGAAATGGGCTATCCCAATCCC 1040
Db 1032 AAGTAAATTTGAAACAGGAAGATGCTGCCAGCTTAAATGAAATGGGACCTCCCAAGGAA 1091
Qy 1041 ATGCAATGTGATGCTACCACTGTGCTGCTGACATTTTCAACGCCCGAATGTAGTTAACT 1100
Db 1092 A-----TGGCTCTGCCACTGTTGCTGACATTTTCACTCCCAAGATGTGCTCAGCT 1142
Qy 1101 TGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTG 1160
Db 1143 TGAGCAGGCTCGGATTTGAGAAATTTGGCAAAATTTAGAGAACCGAAAGAACACTGGCTTG 1202
Qy 1161 TTGTGCTTTATGACCACTGTCCTTACTGCCAGCTATGAGGAAATCTTATGTTGACT 1220
Db 1203 TTGTGCTTTATGACCACTGTCCTTACTGCCAGCTATGAGGAAATCTTATGTTGACT 1262
Qy 1221 TAGCAGCAAGTTAGCAGGCTCAACAGGATGAAGTTGAGAAATTTAGAGCAGATGGAG 1280
Db 1263 TGGCAGAGAAATTTAGCAAGGT---CAGGAGTGAAGTTGCAAAATTTAGAGCAGATGGAG 1319
Qy 1281 AACAGAAAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACCATATTTT 1340
Db 1320 AGCAGAAAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACCATATTTT 1379
Qy 1341 TCCCAAGAGCTTCTCTCGGCGCAACATAAAGTATCCCTCAGAAAAGAGAGATGTTGATT 1400
Db 1380 TCCCAAGAGCTTCTCT---CAACAAATTAAGTACCCTTCAGAAAAGAGAGATGTTGATT 1436
Qy 1401 CTTTGATGGCAATTTGAAATGCTTAAAGTGAAGAT 1436
Db 1437 CATTGACGGCAATTCGTGAATGCTTACGGTGAATG 1472

RESULT 3

US-09-720-318A-9
; Sequence 9, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A

; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-318A-9

Query Match 33.0%; Score 538.2; DB 3; Length 1827;
Best Local Similarity 68.1%; Pred. No. 5,7e-152;
Matches 801; Conservative 0; Mismatches 358; Indels 18; Gaps 3;

Qy 257 GCTTCTGAGACGAAAGGAAGATTTTGAACAGATAGCCAGTGCATCTCGCAATGCTTCA 316
Db 323 GCATCTGCCGTGGCGCGCTGGACTAGAGGCCCTGGCGAGGAGCTGTGGGCGCGTCG 382
Qy 317 CTTCTTGAATCATATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTC 376
Db 383 CGCTGGAGATCATGATCGTGGCTCGACATGTTGGCTCCGAAATCGCCATCGCCTTC 442
Qy 377 AGTGTGCTGAAGATGTTGCTTTGATGATGATGCGAAATTTGACGGTTCGACCTTTAGG 436
Db 443 AGTGTGCTGGAGACGCTGGCCCTCATCGAATACGCAAACTGACTGCGACGCCCTTCAG 502
Qy 437 GTTTTCAGTTTGGACACTGGGAGACTGMAACCCAGAACTTATCAACTTTTGTATGCGTT 496
Db 503 GTTTTCAGCTTGGACACTGGGAGACTGMAACCCAGAACTTATCAACTTTTGTATGCGTT 562
Qy 497 GAGAACATTTATGGAATTCGCATTGATGATCATGTTCCCTGATGCTGTGAGGTTTCAGGCA 556
Db 563 GAGAACATTTATGGAATTCGCATTGATGATCATGTTCCCTGATGCTGTGAGGTTTCAGGCA 622
Qy 557 TTGCTGAGGATGAAGGGTTATTTCTTTCTACGAGATGGGCAACGAAGTGTGTCAGG 616
Db 623 CTTGTGAGGAGCAAGGGCTCTTCTCTTCTACGAGACGGAACACAGGAGTGTGTCAGG 682
Qy 617 GTGAGAAAGGTGAGGCTTTAAGGAGGCGCTTAAAGGCTCTAGAGCATGGAATTAACGTT 676
Db 683 GTGAGAAAGGTGAGGCTTTAAGGAGGCGCTTAAAGGCTCTAGAGCATGGAATTAACGTT 742
Qy 677 CAGAGAAAGACCACTGCTGATGATGCTGTAATACCGGTTGTTTCAAGTTGATCCG 736
Db 743 CAGAGAAAGACCACTGCTGATGATGCTGTAATACCGGTTGTTTCAAGTTGATCCG 802
Qy 737 GCTTTTGAAGGATGATGTTGAAATTTGAAAGCTTGGTGAAGTGGAAACCTGTTGCAAT 796
Db 803 TCTTTTGAAGGATGATGTTGAAATTTGAAAGCTTGGTGAAGTGGAAACCTGTTGCAAT 862
Qy 797 GTGAGGCGCATGACATATGGAATTTCTTCTAGGACCATGATGCTGCTGTAATTCCTTG 856
Db 863 GTGAGGCGCATGACATATGGAATTTCTTCTAGGACCATGATGCTGCTGTAATTCCTTG 922
Qy 857 CATGCAAAAGGATGATGTTTCCATTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGG 916
Db 923 CATGCTCAAGGCTGATGTTTCCATTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGG 982
Qy 917 CAACATGAAGGAGGAGGAGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGTCCTT 976
Db 983 CAGCAGAGAGGAGGAGGAGTGGTGGGAGGAGCGCCACCAAGGAGTGGGCGCTG 1042
Qy 977 CACAAAGGAAATGTAAGCAGCAAGAGGAGGATGTTAATGGAATGGGCTATCCAA 1036
Db 1043 CACAAAGGAAATGTAAGCAGCAAGAGGAGGATGTTAATGGAATGGGCTATCCAA 1090
Qy 1037 TCCCATGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
Db 1091 AAGGCAACGGCTCGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150
Qy 1097 AACTTGAGCAGGATCGAATTTGGAATTTGGCAAAATTTGAGGACCGGAAAGGAAACCATGG 1156

Db 1151 AACCTACCCGTCGGGATCGAGAACTCTCTGGGCTCGAGAACCGCGCGGCGTGG 1210
Qy 1157 CTTGTTGCTTTATGACCAATGTTGCCCCCTACTGCGCAGGCTATGAGGAATCTTATGTT 1216
Db 1211 CTCACCGTCTCTACGCTCCCTGCTGCCCCACTACTGCGCAGGCAATGAGGGCTCTCTACGTT 1270
Qy 1217 GACTTAGCAGACAAGTTAGCAGGCTCAACAGGATGAAGTTGGAATTTAGAGCAGAT 1276
Db 1271 GAGCTGCGCGAAGCTGAGCGGCT---CAGGCATCAAGGTGGCCAAAGTTCCGCGCGGAC 1327
Qy 1277 GGAGAACAGAAAGAAATTTGCAAGAGTGAACTGCAATTTGGGAAGCTTCCCTTACGATATTA 1336
Db 1328 GCGCAGCAGAACCATTCGCGCAGCGGAGCTGCAACTACAGAGCTTCCCGACGATCCTC 1387
Qy 1337 TTTTCCCAAGCAATGCTCTCGGCGCAACAATAAGATATCCCTCAGAAAGAGAGATGTT 1396
Db 1388 CTGTTCCCGCGCCGACCGTGAAGC---CCATCAAGTACCGCTCCGAGAGAGGAGCGTC 1444
Qy 1397 GATTCCCTTGATGGCATTTGTAATGCCCTTAAGATGAG 1433
Db 1445 CAGTCCCTCTCGGCTTCGTGAACAGCCTCAGATGAG 1481

RESULT 4

US-09-720-318A-3

; Sequence 3, Application US/09720318A

; Patent No. 6730827

; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Allen, Stephen M.

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-C

; CURRENT APPLICATION NUMBER: US/09/720,318A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 3

; LENGTH: 1210

; TYPE: DNA

; ORGANISM: Impatiens balsamia

US-09-720-318A-3

Query Match 26.7%; Score 435.2; DB 3; Length 1210;

Best Local Similarity 69.7%; Pred. No. 6.5e-121;

Matches 641; Conservative 0; Mismatches 258; Indels 21; Gaps 3;

Qy 516 GCATTGAGTACATGTTCCCTCATGCTGTTGAGGTTTCAGGCAATGTTAGTCAAGGATTTAGTAAGAACCCAAAGGAC 575
Db 1 GCACGAGGTACATGTTCCCTCATGTCATTTGTTAGTCAAGGATTTAGTAAGAACCCAAAGGAC 60
Qy 576 TATTCTCTTTCTACGAGGATGGCCACCAAGAGTGTTCAGGCTTCAGGCTGAGAAAGTGAAGGCTT 635
Db 61 TGTCTCTTTCTACGAGAGCGACATCAAGAGTGTGCGCGCTCAGAAAGTGAAGGAC 120
Qy 636 TAAGAGGGCCCTTAAGGCTCTCAGAGCATGGAATCTGTTGTCAGAGGAAAGACCAAGTCAAC 695
Db 121 TGAGGCGTGTCTCAAGGCTCTCCGCGCTTCGATCAGCGGGCAAGAAAGACCAAGTCCG 180
Qy 696 CTGGTACTAGTCTGAATACCGTGTTCAGGTTGATCCGGCTTTTCAGGGAATGGATG 755
Db 181 CGGGAAACAGATCGGAGATCCAGTTCGTCAGTGAATCCCTCTTTTGAAGGATTTGGTTG 240
Qy 756 GTGGAATTTGGAAGCTTGTGTAAGTGGAAACCCCTGTTTGCATAATGTTGAAGGCCCATGACATAT 815
Db 241 GTGGAGAGGTTAGCTGTGTGAAGTGAATCCGCTGCTTAATGTAGATGCTGCTGATGAT 300
Qy 816 GGAATCTCTTAGGACCAATGAATGCTGCTGGAATTCCTTGATGCCAAAGAGATATGTT 875
Db 301 GGAATTTCTCGAGCTATGAATGCTGCTGTTAATGCACCTTCATAGCAGGAGGTTATGCT 360
Qy 876 CCATTGGGTGTGAGCCCTGCATAGGCTGTTTACCTGGGCAACATGAAGGGAAGGA 935

Db 361 CGATTGGTGGACCGTGCACCCGCGGTGTACTCTGGCAACATGAGAGAGAGCA 420
Qy 936 GGTGGTGGTGGAGGATGCGCAAGCTAAGGAATGTGTCTTTCACAAAGGAATTAAGC 995
Db 421 GGTGGTGGTGGAGGATGCTGCGGCTAAGGAGTGTGCGCTTACATAAAGGAATATAA--- 477
Qy 996 AGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCCAAATGGTGATG 1055
Db 478 -----AGGATGCCAATGGGAATGGGTTGCTCAAGCTGAGGGAGGAAGAA 525
Qy 1056 CTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAGGACTGGA 1115
Db 526 CTGTTACGGATGCTGATATTTTGAATCCAAGAAATGTGTGACACTCAGTAGAAGCGGA 585
Qy 1116 TTGAGAATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGTGCTTTATGCAC 1175
Db 586 TTGAGAATCTGTGNAACTTTCAGGAGAGGAAGAGCCATGGATCGTGTCTGTATGCAC 645
Qy 1176 CATGGTCCCTACTGCGCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGACAAGTTAG 1235
Db 646 CTTGGTCCAGTCTCTGCCAGGATATGGAATAATCATACTTGGAAATTTGCTGAAAGCTGG 705
Qy 1236 CAG---GGTCAACAGGATGAAGGTTGGAATAATTTAGAGCAGATGGAGAACAGAAAGAT 1292
Db 706 CGGTGAGCGGTGGTGTGAGAGGTAGGGAATTTCCGGCAGATGCTGCAGAAAGGAGT 765
Qy 1293 TTGCAAGAGTGAACTGCAATTTGGGAAGCTTCCCTCAGATATTAATTTTCCCAAGCAAT 1352
Db 766 TTGCTACCAAGAAATTCAGCTGGGAGCTTTCACAACTACTCTTCTTCCCAACACT 825
Qy 1353 CGTCTCGGCCCAACAAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCAT 1412
Db 826 CATCT---AAAGCCATCAAGTACCATCTGAGAAAAAGGAGCGTGGAGTCATTGTTGGCTT 882
Qy 1413 TTGTAATGCTTAAAGATGA 1432
Db 883 TTGTGAACGCACTCAGATGA 902

RESULT 5

US-09-720-318A-1

; Sequence 1, Application US/09720318A

; Patent No. 6730827

; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Allen, Stephen M.

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-C

; CURRENT APPLICATION NUMBER: US/09/720,318A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1215

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (273)

; OTHER INFORMATION: n = a, c, g or t

US-09-720-318A-1

Query Match 22.4%; Score 364.2; DB 3; Length 1215;

Best Local Similarity 65.2%; Pred. No. 2e-99;

Matches 591; Conservative 0; Mismatches 294; Indels 21; Gaps 3;

Qy 528 TGTTCCTGATGCTGTGAGGTTTCAGGCAATGTTGAGGAGTAAGGGTTATCTTTCT 587
Db 1 TGTTCCTGAGCGCGAGGCTGAGGAGTGGTGGCGCACCAAGGGCTCTTCTCTCTCT 60


```
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13944

Query Match      9.2%; Score 149.4; DB 3; Length 810;
Best Local Similarity 52.2%; Pred. No. 1.6e-34;
Matches 361; Conservative 0; Mismatches 321; Indels 9; Gaps 1;

Qy 302 CTGACAAATGCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAAGCAC 361
Db 115 CTTGCCGACAAATCCCCCGAGGACATCCTGAAAGCCGCTTCGAGACATCTTCGGCGAG 174
Qy 362 ATAGCTATTGCTTCAGTGTGCTGAAGATGTTGCTTTGATTGATGATGCGAAATTCAGC 421
Db 175 CTGTGGATCTCTTCACGGCGCGAGGAGCGTGTCTAGTAGACATGGCTGGAGCTC 234
Qy 422 GGTGACCCCTTTAGGGTTCCTGAGTTGGACACTGGGAGACTGAACCCAGAAATTCATCAA 481
Db 235 AACCGCAACGTGAAGGTGTTTCAGCTCGACACCGGTGCGCTGCACCCGGAACCTATCGT 294
Qy 482 CTTTTCATGCGGTGAGAGCATATGGAATTCGCAATTCGATGATGATGATGATGATGATG 541
Db 295 TTATATGACACGAGTTCGCGCAACTACGCGATCCGATCGACGTGCTTCCCGCGACCCG 354
Qy 542 GTTGAGGTTCAGGCAATTCGAGGAGTAAAGGGTTCCTTTCTTCTACGAGGATGGGCAC 601
Db 355 CGCTGCTCGAACCCCTGTTGAGGAAAGGGCCCTGTTACGCTTCTACCGGAGCGGCAC 414
Qy 602 CAAGAGTGTTCAGGGTGAAGAGGTGAGGCTTTTAAAGGAGGGCCCTTAAAGGCTTCAGA 661
Db 415 GGTGAGTGTTCGCGCATCCGCAAGATCGAACCGCTCAAGCGCAAGCTCGCGCGGTGCGC 474
Qy 662 GCATGATTAATGCTGACAGGAAGACAGTCACTGCTAGTACTAGTCTGAAATACCGTT 721
Db 475 GCCTGGGCGCACCGGCCAACCGCGCACGAGAGCCCGCGCACGCGAGCGGTGGCGGTG 534
Qy 722 GTTCAGGTTCATCGGCTTTTGGGGAATGATGATGATGATGATGATGATGATGATGATG 781
Db 535 CTGGAATTCAGCGTGCCTTTTCCAGCGCGA-----AAAGCGCTGTACAAATTC 585
Qy 782 AACCTGTGTGCAATGTGAAGGGCCATGACATATGGAATTCCTTTAGGACCATGAATGTG 841
Db 586 AACCCACTGCTGCTCATGATCAGCAGGAGGCTCTGGGGCTATATCCGATGCTCGAACTG 645
Qy 842 CCTGTGAATTCCTTGCATGCAAAAGGATATGTTCCATTTGGTGTGAGCCCTGCACATAG 901
Db 646 CCCTACACAGCTGACGAAACGCGCTATATCAGCATCGGCTCGGAACCCCTGCACCCGT 705
Qy 902 CTTGTTTACCTGGCAACATGAAGGAGGAGGTGTTGTTGGGAGGATGCCAAAGCT 961
Db 706 CCGTGTCTGCCAACACGACGAGCGGAGGCGCGCTGTTGTTGGGAGGAGCCACCCAC 765
Qy 962 AAGGAATGTGCTTCTTCAAAAGGAATGTAA 992
Db 766 AAGGAGTGGGGCTGCACCGCGCAACCTGA 796

RESULT 8
US-09-252-991A-13837
; Sequence 13837, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 13837
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13837

Query Match      5.2%; Score 84.2; DB 3; Length 660;
Best Local Similarity 54.8%; Pred. No. 7.4e-15;
Matches 167; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 302 CTGACAAATGCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAAGCAC 361
Db 356 CTTGCCGACAAATCCCCCGAGGACATCCTGAAAGCCGCTTCGAGACATCTTCGGCGAG 415
Qy 362 ATAGCTATTGCTTCAGTGTGCTGAAGATGTTGCTTTGATTGATGATGCGAAATTCAGC 421
Db 416 CTGTGGATCTCTTCAGCGCGCGCGGAGACGTGCTCTAGTAGACATGGCTGGAGCTC 475
Qy 422 GGTGACCCCTTTAGGGTTCCTGAGTTGGACACTGGGAGACTGAACCCAGAAATTCATCAA 481
Db 476 AACCGCAACGTGAAGGTGTTTCAGCTCGACACCGGTGCGCTGCACCCGGAACCTATCGT 535
Qy 482 CTTTTCATGCGGTGAGAGCATATGGAATTCGCAATTCGATGATGATGATGATGATGATG 541
Db 536 TTATATGACACGAGTTCGCGCAACTACGCGATCCGATCGACGTGCTTCCCGCGACCCG 595
Qy 542 GTTGAGGTTCAGGCAATTCGAGGAGTAAAGGGTTCCTTTCTTCTACGAGGATGGGCAC 601
Db 596 CGCTGCTCGAACCCCTGTTGGAAGAAAGGGCTGTTACGCTTCTACCGGAGCGGCAC 655
Qy 602 CAAGA 606
Db 656 GGTGA 660

RESULT 9
US-09-902-540-7978
; Sequence 7978, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7978
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7978

Query Match      5.0%; Score 81; DB 3; Length 723;
Best Local Similarity 49.6%; Pred. No. 7.3e-14;
Matches 346; Conservative 0; Mismatches 315; Indels 36; Gaps 4;

Qy 292 AGCCAGTGTATCTCGACAAATGCTTCACCTCTTGAATCATGATAGAGCCCTCGACAAAT 351
Db 42 AGCCGCTGAGTGAAGAGCGCCCGCGAGACCTCTCGCTGGACCGAGCGCGCTT 101
Qy 352 CGGCAACGACATAGCTATTGCTTCAG---TGGTGTGGAAGATGTTGTTGATTGATGATA 408
Db 102 CGGTGCGCGCGCGGCCCATCGCTCCAGCTTCGGGCTGGAGGACATGTTCTCATCGACT 161
Qy 409 TGGCAATTCAGCGGTGCGACCCCTTTAGGTTTTCAGTTTTCAGACTGGAGACTGAACCC 468
Db 162 GCGCGCGCAGCATCGGCCAGCTTCGCGCTCTTTCACGCTCGACCCGAGCGCTGCCCCC 221
```

QY 469 AGAACTTATCACTTTTGTGATGCGTTGAGAGCAATATGGAATTCGCATGAGTACAT 528
Db 222 GGAGAGCTTACAACTCATGAGGTGGTGGCGTACCGGCTCACCGTGGAGACGTA 281
QY 529 GTTCCCTGATGCTGTTGAGGTTTCAAGGATGAGGAGTAAGGGGTTATCTCTTTCTA 588
Db 282 CTTCCCGGAGCGCGCGGTGGAGGCTTGGAGTCCAGAACGGCTACTTCTCTTCCG 341
QY 589 CGAG-----GATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGTGAGGCTTTTAAGGAG 642
Db 342 CCAGAGCTAGAGCGCAGCAAGCGCTGCTGGCCATCCGCAAGTGGAGCCCTGTGCG 401
QY 643 GGCCCTTAAGGGTCTCAGAGCATGATGTAATCTGTGAGAGAAAGACAGTCACTGGTAC 702
Db 402 CGCGCTCGCGGGTCAACAGAGCGTGGGTGACGGATTCGCGCTGAGCAGT---CCGTAC 458
QY 703 TAGGTCTGAATACCGGTTGTTGAGGTTGATCCGGCTTTTGAGGGATGATGGTGAAT 762
Db 459 CCAGACCGAGCTGGCAGCTTAGAGGTGGACAGCGGCATG----- 499
QY 763 TGAAGCTTGGTGAAGTGAACCTGTTGCAATGTGAAGGGCCATGACATATGGAATTT 822
Db 500 -----GGCTGCTCAAGCTCAACCGCTGGCCAGCTGGAGCAGCGGGACATCTGGGCCTA 554
QY 823 CTTAGAACCATGAATGTGCTGTGAATCTCTTGCATGCAAAAGGATATGTTTCCATTTG 882
Db 555 CGTCCGGGCGAAGAGCGTGCCTTACAAACGCGCTCCATGACCGTGGCTACCCGCTCCATCG 614
QY 883 GTGTGAGCCTGCACTAGGCTGTTTACCTGGGCAACATGAAGGAGGAGGAGTGGTG 942
Db 615 CTGTGCGCCCTGACGCGCGCTGAAACCTTACGAGGACGAGCGCGCGCGCTGGTG 674
QY 943 GTGGAGGATGCCAAAGCTAAGGAATGTGCTTTTCC 979
Db 675 GTGGAGTCCGCGGAGAACCGGAGTGGCGGCTCCAC 711

RESULT 10

US-09-902-540-807/c
; Sequence 807, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 807
; LENGTH: 6063
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-807

Query Match 5.0%; Score 81; DB 3; Length 6063;
Best Local Similarity 49.6%; Pred. No. 2.5e-13;
Matches 346; Conservative 0; Mismatches 315; Indels 36; Gaps 4;
QY 292 AGCCAGTGATCTCGACAATGCTTTCACCTCTTGAATCATGATAGAGCCCTCGACAAAT 351
Db 3803 AGCCGCTGAAGTGAAGAGAGCCCGCGGAGGACCTCTCGCTGACGAGCGCGCTT 3744
QY 352 CGGCAACGACATAGCTATTGCTTTTCAG---TGCTGTGAAGATGTTGCTTTTGAATGAGTA 408
Db 3743 CGGTGCGCGCGCGCCATCGCTTCCAGCTTCGCGGTGGAGGACATGCTCTCATCGACCT 3684
QY 409 TGGAAATGACGGGTGACCCCTTTTAGGTTTTCAGTTTGGACACATGGGAGACTGAACCC 468

Db 3683 GGCGCCCGACGATGCGCCCGAGCGCTGCGCTTTTACGCTGACACCGGACGCGCTGCCCC 3624
QY 469 AGAACTTATCACTTTTGTGATGCGGTTGAGAGCAATATGGAATTCGCATGAGTACAT 528
Db 3623 GGAGAGCTTACAACTCATGAGGTGGTGGCGTACCGGCTCACCGTGGAGACGTA 3564
QY 529 GTTCCCTGATGCTGTTGAGGTTTCAAGGATGAGGAGTAAGGGGTTATCTCTTTCTA 588
Db 3563 CTTCCCGGAGCGCGCGGTGGAGGCTTGGAGTCCAGAACGGCTACTTCTCTTCCG 3504
QY 589 CGAG-----GATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGTGAGGCTTTTAAGGAG 642
Db 3503 CCAGAGCTTAGAGGCGCAGCAAGCGCTGCTGGCCATCCGCAAGTGGAGCCCTGTGCG 3444
QY 643 GGCCCTTAAGGGTCTCAGAGCATGGAATACTGTGAGAGAAAGACAGTCACTGGTAC 702
Db 3443 CGCGCTCGCGGGTCAACAGCGCTGGGTGACGGATTCGCGCTGAGCAGT---CCGTAC 3397
QY 703 TAGGTCTGAATACCGGTTGTTGAGGTTGATCCGGCTTTTGAGGGATGATGGTGAAT 762
Db 3386 CGCAGCGAGCTGGCGAGCTTAGAGGTGGACAGCGGCATG----- 3346
QY 763 TGAAGCTTGGTGAAGTGAACCTGTTGCAATGTGAAGGGCCATGACATATGGAATTT 822
Db 3345 -----GGCTGCTCAAGCTCAACCGCTGGCCAGCTGGAGCAGCGGGACATCTGGGCCTA 3291
QY 823 CTTAGAACCATGAATGTGCTGTGAATCTTTCATGCAAAAGGATATGTTTCCATTTG 882
Db 3290 CGTCCGGGCGAAGAGCGTGCCTTACAAAGGCTCCATGACCGTGGCTACCCGTTCCATCG 3231
QY 883 GTGTGAGCCTGCACTAGGCTGTTTTCATCTGGGCAACATGAAGGAGGAGGAGTGGTG 942
Db 3230 CTGTGCGCCCTGACGCGCGCTGAAACCTTACGAGGACGAGCGCGCGCGCTGGTG 3171
QY 943 GTGGAGGATGCCAAAGCTAAGGAATGTGCTTTTCC 979
Db 3170 GTGGAGTCCGCGGAGAACCGGAGTGGCGGCTCCAC 3134

RESULT 11

US-09-596-002-41
; Sequence 41, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 269223
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 41
; PUBLICATION INFORMATION:
US-09-596-002-41

Query Match 4.5%; Score 73.4; DB 3; Length 269223;
Best Local Similarity 58.4%; Pred. No. 4.5e-10;
Matches 128; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 771 TGTGAGTGGAGAACCTGTTGCAAAATGTGAAGGCCCATGACATATGAACTTCTTAGGA 830
Db 233373 TTGCCAAATACAAATCCAAATTTTGTATGGCAGAGACGATGTATGGCATATATTTGA 233432

```
Qy 831 CCATGATGCGCTGCTGAATTCCTTGATGCAAGAGGATATGTTCCATTGGGTGAGC 890
Db 233433 CCAAAATATACCGTTTAATGAGCTGTATCACAAGGCTATCCTTCCATTGGCTGTAGC 233492
Qy 891 CCTGCACTAGGCTGTTTACCTGGGCAACATGAAAGGAGGAGGTGCTGTTGGGAGG 950
Db 233493 CTTGTACCATCGCTGTCAAGCAGGCGAGGATATCCGTGCAGGACGCTGCTGGTGGGAGC 233552
Qy 951 ATGCCAAGCTAAGGAATGCTGCTTCCACAAAGGAATG 989
Db 233553 ATAAGATAATAAGAGTGGGTTTACACAATGACTTG 233591

RESULT 12
US-09-540-236-1037
; Sequence 1037, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1037
; LENGTH: 756
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-1037

Query Match 4.5%; Score 72.6; DB 3; Length 756;
Best Local Similarity 58.6%; Pred. No. 2.6e-11;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 771 TGGTGAAGTGGAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGGAATCTTCCTTAGGA 830
Db 542 TTGCAAAATACATCCAAATTTGATTTGGCAGAGAGGATGATGAGCATATATTGGA 601
Qy 831 CCATGAATGTGCTGTGAATTCCTTGATGCAAAAGGATATGTTTCCATTGGGTGAGC 890
Db 602 CCAAAATATACCGCTTTAATGAGCTGTATCACAAGGCTATCCTTCCATTGGCTGTGAGC 661
Qy 891 CCTGCACTAGGCTGTTTACCTGGGCAACATGAAAGGAGGAGGTGCTGGTGGGAGG 950
Db 662 CTTGTACCATGCTGTCAAGCAAGCGGAGGATATCCGTGCAGGACGCTGTGGTGGGAGC 721
Qy 951 ATGCCAAGCTAAGGAATGCTGCTTCCACAAAGGA 985
Db 722 ATAAGATAATAAGAGTGGCGTTTACACAATGA 756

RESULT 13
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOVIPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 435
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 3.6%; Score 59; DB 2; Length 7218;
Best Local Similarity 1.3%; Pred. No. 1.3e-06;
Matches 5; Conservative 230; Mismatches 140; Indels 0; Gaps 0;

Qy 924 AAAGGAGGAGGAGGTGCTGGGAGGATGCCAAGCTAAGGATGTGGTCTTCACAAAG 983
Db 1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
Qy 984 GAAATGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATG 1043
Db 1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
Qy 1044 CAAATGTGTGTACCATGCTGCTGACATTTTCAACAGCCCGAATGTAGTTAATTGA 1103
Db 1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
Qy 1104 GCAGGACTGGAATTGACAATTTGCCAAATTTGGAGGACCGAAGAACCATGCTGTGTG 1163
Db 1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
Qy 1164 TGTCTTATGCACCATGCTGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTACTAG 1223
Db 1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
Qy 1224 CAGACAAAGTTAGCAGGCTCAACAGGATGAAGTTGGAAATTTAGACGACATGGAGAAC 1283
Db 1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
Qy 1284 AGAAAGAATTTGCAA 1298
Db 1074 RRRRRRRRATGCGCAA 1060

RESULT 14
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
```



```
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match      2.6%; Score 42; DB 3; Length 1141;
Best Local Similarity 9.8%; Pred. No. 0.06;
Matches 45; Conservative 214; Mismatches 191; Indels 10; Gaps 2;

Qy 1153 ATGCTTGTGCTTATGACCATCGTCCCTACTGCCAGGCTATGAGGAACTTA 1212
Db 159 DKTRTWWKNNNATGDDDTKYHMMNNNGCBVTVMVRYKTRDWSBKRMNYGMBWKN 218
Qy 1213 TGTGACTTAGCAGACAAAGTTAGCAGGTCACAGGATGAAGTTGGAATAATTAGAGC 1272
Db 219 WSDVDTYYWVDDMCCKYRWRVTRGRMNYVAVBTAHRRYNNGWTEBAYRWTM 278
Qy 1273 AGATGGAGACAGA-----AAGAAATTGCAAGAGTGAACATGCAATTGGGAAGCTT 1323
Db 279 NNNNNNAKMKRAKYGNWRABVNSTCTTKWSKTKTKVTSWANNCRAGDANKDHKKWK 338
Qy 1324 CCTACATATATTTTCCCAAGCAATTCGTCTCGGCCCAACAATAAGTATCCTCAGA 1383
Db 339 WSAAMGYWNNNNNNWNTYKARHBAWHDVWHSAMWCKHANAHAHYSRKKWTBYKRTMV 398
Qy 1384 AAAGAGAGATGTTGATTCCTTGATGCGATTTGTAATGCCTTAAGATGAGGATATCAGGA 1443
Db 399 NNNNGTTWKBMAWYKMDMDHBTGTYNNNNGRTYTGTKKMTYKWKANNCKWR 458
Qy 1444 AATTTCTTCGTTTGGGTGCAATTCACATTTGACTACGTACAGCGGGTTCCTTCT 1503
Db 459 AWDHKTCTHNTTWMKMTYNNNCKYKSMNTNGSKSHRBAAAVYTWYMMWRRYAHANNW 518
Qy 1504 TTATGCTATGCTGATATACCATTCGTTTACAGATTTCTTGCAACTCGTTGGAAG 1562
Db 519 DYWKACTYKYBVCCKMNNYAAYTKSSWNTYSRYRWKTNNSWRWSRSDTRSMGRANN 578
Qy 1563 TGGGAATGGAGGTTTATACAAATAAGATACCTCAGTTTGA 1602
Db 579 YARABHYGKWNTRWBSHWTBHBRAAGAAHYWBMWYBA 618

RESULT 15
US-09-385-982-465/c
; Sequence 465, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)..(578)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-465
```

```
Query Match      2.4%; Score 39.2; DB 3; Length 578;
Best Local Similarity 48.0%; Pred. No. 0.28;
Matches 110; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 1401 CCTTGATGGCATTTGTAAATGCCTTAAGATGAGGATATCAGGAAATTTCTTCGTTTGG 1460
Db 243 CCTTCTTAGAATGTGCAAGTTCCATGAGAGTAATATATGATGATTTTGAAGAAGATT 184
Qy 1461 GGTGCAATTCACATTTTGACTATACGTACAGCGGGTTCCTTCTTTATGCTATTACGTGTA 1520
Db 183 GTTGATAGTTACATCTTCAAACTTATCATTTCCAGTATGCACTTTTAAGATAATGTGATTC 124
Qy 1521 TATACCATTCGTTTACAGATTTCTTCTGTGAACCTCGTTGGAAGTGGGAATGGAGTTTATA 1580
Db 123 TAAAGTAGATGACTTTATATTTCTTGTATTAAGAGTGTCTATACATGTTAAGAAATGCAATAA 64
Qy 1581 CAAATAAGATACCTCAGTTTGTGAATGTTTTTAAAAAATTTTTTTTTTTTTTTTTTTT 1629
Db 63 GGAATACAAATAAATATTTTAAAGTGTATGTAAGAAAAAATTTTTTTTTTTTTTTTTTTT 15
```

Search completed: February 20, 2006, 23:21:19
Job time : 320 secs

LEFT BLANK

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 21:15:06 ; Search time 1412 Seconds
(without alignments)
9540.239 Million cell updates/sec

Title: US-10-731-525-7
Perfect score: 1629
Sequence: 1 gcacggaggagaccata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1629	100.0	1629	7	US-10-731-525-7
2	1629	100.0	1629	7	US-10-762-049-7
3	916.8	56.3	1730	7	US-10-425-114-7777
4	916.8	56.3	2774	7	US-10-424-599-134742
5	904	55.5	1795	7	US-10-731-525-5
6	904	55.5	1795	7	US-10-762-049-5
7	685.6	42.1	711	7	US-10-424-599-135400
8	651.6	40.0	1398	3	US-09-938-842A-498
9	651.6	40.0	1398	3	US-09-938-842A-498
10	620.6	38.1	1796	7	US-10-424-599-82919
11	577.4	35.4	1365	3	US-09-938-842A-2305
12	577.4	35.4	1365	3	US-09-938-842A-2305
13	538.2	33.0	1827	7	US-10-731-525-9
14	538.2	33.0	1827	7	US-10-762-049-9
15	512.6	31.5	1840	8	US-10-739-930-2784
16	472.4	29.0	1553	8	US-10-425-115-83222
17	472	29.0	1880	7	US-10-437-963-23751
18	467.4	28.7	1380	7	US-10-169-667A-1
19	435.2	26.7	1210	7	US-10-731-525-3
20	435.2	26.7	1210	7	US-10-762-049-3
21	425.4	26.1	1850	8	US-10-425-115-83219
22	419.6	25.8	1927	8	US-10-425-115-83223
23	369.4	22.7	549	8	US-10-425-115-147208

24	364.4	22.4	1156	7	US-10-425-114-12185	Sequence 12185, A
25	364.2	22.4	1215	7	US-10-731-525-1	Sequence 1, Appli
26	364.2	22.4	1215	7	US-10-762-049-1	Sequence 1, Appli
27	356.6	21.9	1183	7	US-10-425-114-15300	Sequence 15300, A
28	352.8	21.7	1194	8	US-10-425-115-83221	Sequence 83221, A
29	277.4	17.0	623	7	US-10-021-323-11704	Sequence 11704, A
30	266.6	16.4	471	3	US-09-732-627A-2057	Sequence 2057, Ap
31	265.6	16.3	1294	8	US-10-425-115-83224	Sequence 83224, A
32	256.4	15.7	585	7	US-10-021-323-3831	Sequence 3831, Ap
33	232.4	14.3	985	7	US-10-767-701-10835	Sequence 10835, A
34	223.8	13.7	588	7	US-10-021-323-3903	Sequence 3903, Ap
35	211.6	13.0	552	7	US-10-021-323-11747	Sequence 11747, A
36	186	11.4	330	7	US-10-767-701-17535	Sequence 17535, A
37	161.8	9.9	599	7	US-10-424-599-69970	Sequence 69970, A
38	159	9.8	705	6	US-10-369-493-37482	Sequence 37482, A
39	141.4	8.7	514	3	US-09-770-961-142	Sequence 142, App
40	132.4	8.1	742	8	US-10-425-115-12975	Sequence 12975, A
41	119.4	7.3	453	3	US-09-732-627A-2083	Sequence 2083, Ap
42	117.8	7.2	492	8	US-10-767-795-6146	Sequence 6146, Ap
43	109.4	6.7	667	8	US-10-425-115-156122	Sequence 156122, A
44	95.8	5.9	213	7	US-10-424-599-4794	Sequence 4794, Ap
45	81	5.0	690	6	US-10-369-493-42894	Sequence 42894, A

ALIGNMENTS

RESULT 1
US-10-731-525-7
; Sequence 7, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10731,525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-10-731-525-7

Query Match	100.0%	Score 1629;	DB 7;	Length 1629;
Best Local Similarity	100.0%	Pred No. 0;		
Matches 1629;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCACGAGGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTTCACCTTCAATTC	60	
DB	1	GCACGAGGAGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTTCACCTTCAATTC	60	
QY	61	CGCACCACTTCCACCTTCCCATCATCGGACCCAACTCCGCAATTCGGTCAATAG	120	
DB	61	CGCACCACTTCCACCTTCCCATCATCGGACCCAACTTCGCAATTCGGTCAATAG	120	
QY	121	GATTCGGAGAGGCCCATTTGGAGCGCCGTTAATTTCAATTTATCTCAAGAGCGAGCTT	180	
DB	121	GATTCGGAGAGGCCCATTTGGAGCGCCGTTAATTTCAATTTATCTCAAGAGCGAGCTT	180	
QY	181	GGTAAAGCCGTTAACCGCAACCTCCACGCAAGGATTCATTTGCTCTCCGAGCAAC	240	
DB	181	GGTAAAGCCGTTAACCGCAACCTCCACGCAAGGATTCATTTGCTCTCCGAGCAAC	240	
QY	241	AACCATCGTGTCTTCTGCTCTGAGACGAAAGGAGATTGTAACAGATAGCAGTGA	300	
DB	241	AACCATCGTGTCTTCTGCTCTGAGACGAAAGGAGATTGTAACAGATAGCAGTGA	300	

```
Qy 301 TCTCGAATGCTTCACTCTTGAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360
Db 301 TCTCGAATGCTTCACTCTTGAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360
Qy 361 CATAGCTATTGCGCTTCAGTGGTCTGAAGATGTTGCTTTGATTTGAGTATCGAAATTTGAC 420
Db 361 CATAGCTATTGCGCTTCAGTGGTCTGAAGATGTTGCTTTGATTTGAGTATCGAAATTTGAC 420
Qy 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480
Db 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480
Qy 481 ACTTTTGTAGTCGGTTGAGAAAGCATATGAAATTCGCAITTGAGTACATGTTCCCTGATGC 540
Db 481 ACTTTTGTAGTCGGTTGAGAAAGCATATGAAATTCGCAITTGAGTACATGTTCCCTGATGC 540
Qy 541 TGTGTAGGTTTCAGGATTCAGGAGTAAAGGGTTATCTCTTTCTACGAGGATGGCA 600
Db 541 TGTGTAGGTTTCAGGATTCAGGAGTAAAGGGTTATCTCTTTCTACGAGGATGGCA 600
Qy 601 CCAAGAGTGTTCAGGGTGAGAAAGTGAGGCCCTTTAAAGAGGGCCCTTAAAGGCTCTCAG 660
Db 601 CCAAGAGTGTTCAGGGTGAGAAAGTGAGGCCCTTTAAAGAGGGCCCTTAAAGGCTCTCAG 660
Qy 661 AGCATGGATTAAGTGTTCAGAGGAAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGT 720
Db 661 AGCATGGATTAAGTGTTCAGAGGAAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGT 720
Qy 721 TGTTCAGGTTGATCCGGCTTTTGGAGGAATGGATGGTGGAAATGGAGCTTGGTGAAGTG 780
Db 721 TGTTCAGGTTGATCCGGCTTTTGGAGGAATGGATGGTGGAAATGGAGCTTGGTGAAGTG 780
Qy 781 GAACCCCTGTTGCAAAATGTGAAGGCCATGATCATATGCAACTTCTTTAGGACCAATGAATG 840
Db 781 GAACCCCTGTTGCAAAATGTGAAGGCCATGATCATATGCAACTTCTTTAGGACCAATGAATG 840
Qy 841 GCCTGTGAATCTCTGATGCAAAAGGATATGTTCCATTTGGGTGAGCCCTGCACATAG 900
Db 841 GCCTGTGAATCTCTGATGCAAAAGGATATGTTCCATTTGGGTGAGCCCTGCACATAG 900
Qy 901 GCCTGTTTAACTCGGGCAACATGAAGGAGGAGGTGGTGGTGGAGGATGCCAAAGC 960
Db 901 GCCTGTTTAACTCGGGCAACATGAAGGAGGAGGTGGTGGTGGAGGATGCCAAAGC 960
Qy 961 TAAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTAATGG 1020
Db 961 TAAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTAATGG 1020
Qy 1021 AAATGGGCTATCCCAATCCCATGCAATGGTGTATGCTACCACTGTGCTGCATTTTCAA 1080
Db 1021 AAATGGGCTATCCCAATCCCATGCAATGGTGTATGCTACCACTGTGCTGCATTTTCAA 1080
Qy 1081 CAGCCGGAATGTAGTTAACTTGAAGGAGTGGAAATTTGAGAAATTTGGCAAAATTTGAGGA 1140
Db 1081 CAGCCGGAATGTAGTTAACTTGAAGGAGTGGAAATTTGAGAAATTTGGCAAAATTTGAGGA 1140
Qy 1141 CCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGTTGTCCTTACTGCCAGGCTAT 1200
Db 1141 CCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGTTGTCCTTACTGCCAGGCTAT 1200
Qy 1201 GGAGGAATCTTATGTTAGCAGACCAAGTTAGCAGGCTCAACAGGGATGGAAGTTGG 1260
Db 1201 GGAGGAATCTTATGTTAGCAGACCAAGTTAGCAGGCTCAACAGGGATGGAAGTTGG 1260
Qy 1261 AAAATTTAGACAGATGGAGAACAGAAAGATTTGCAAAAGAGTGAATGCAATTTGGGAAG 1320
Db 1261 AAAATTTAGACAGATGGAGAACAGAAAGATTTGCAAAAGAGTGAATGCAATTTGGGAAG 1320
Qy 1321 CTTCCCTACGATATATTTTCCAAAGCATTCGTCTCGGCCCAACAATATAAGTATCCCTC 1380
Db 1321 CTTCCCTACGATATATTTTCCAAAGCATTCGTCTCGGCCCAACAATATAAGTATCCCTC 1380
```

```
Qy 1381 AGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCTTAAGATGAGGATATCA 1440
Db 1381 AGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCTTAAGATGAGGATATCA 1440
Qy 1441 GGAATATTTTCTTCGTTTGGTGGTGCATTTCCATTTGACTATACGTACAGCGGTTTCT 1500
Db 1441 GGAATATTTTCTTCGTTTGGTGGTGCATTTCCATTTGACTATACGTACAGCGGTTTCT 1500
Qy 1501 TCTTTATGCTATTACGTGTATATACCATTCGTTTACAGATTCTTCTGTGAACCTGTTGGA 1560
Db 1501 TCTTTATGCTATTACGTGTATATACCATTCGTTTACAGATTCTTCTGTGAACCTGTTGGA 1560
Qy 1561 AGTGGGAATCGAGGTTTATACAAATAGATACACTCAGTTTGAATGGTTTAAAAAATAA 1620
Db 1561 AGTGGGAATCGAGGTTTATACAAATAGATACACTCAGTTTGAATGGTTTAAAAAATAA 1620
Qy 1621 AAAAAAATAA 1629
Db 1621 AAAAAAATAA 1629

RESULT 2
US-10-762-049-7
; Sequence 7, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCES: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-7

Query Match 100.0%; Score 1629; DB 7; Length 1629;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACTTCTTCAATTC 60
Db 1 GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACTTCTTCAATTC 60
Qy 61 CGCACCAACTTCCACCTTCCCATCATCGGAACCCAAACCTTCGCAAAATGGGTCAATTAG 120
Db 61 CGCACCAACTTCCACCTTCCCATCATCGGAACCCAAACCTTCGCAAAATGGGTCAATTAG 120
Qy 121 GATTTCCGAGAGGCCCATTCGAGCGCGGTAAATTTCAATTTATCTCAAGACCGAGCTT 180
Db 121 GATTTCCGAGAGGCCCATTCGAGCGCGGTAAATTTCAATTTATCTCAAGACCGAGCTT 180
Qy 181 GGTAAAGCCGGTTAAACCCGAAACCTCCACGCAAGGATTCCATTTCTCTCGCAGCAAC 240
Db 181 GGTAAAGCCGGTTAAACCCGAAACCTCCACGCAAGGATTCCATTTCTCTCGCAGCAAC 240
Qy 241 AACCATCGTTGCTTCTGCTTCTGAGAGAAAGAGAAATTTTGAACAGATAGCCAGTGA 300
Db 241 AACCATCGTTGCTTCTGCTTCTGAGAGAAAGAGAAATTTTGAACAGATAGCCAGTGA 300
Qy 301 TCTCGCAATGCTTCACTCTTGAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360
Db 301 TCTCGCAATGCTTCACTCTTGAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360
Qy 361 CATAGCTATTGCTTCACTCTTGAATCATGGATAGAGCTTGTGTTGATGATGCAAAATTCGAC 420
```

```
Db 361 CATAGATATTGCCCTTCAGTGGTCTGAAGATGTTGCTTGAATGAGTATGCGAAATTGAC 420
Qy 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCA 480
Db 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCA 480
Qy 481 ACTTTTGTAGCGGTTCAGAAAGCAATTAAGCAATTCGCAATGAGTACATGTTCCCTGATGC 540
Db 481 ACTTTTGTAGCGGTTCAGAAAGCAATTAAGCAATTCGCAATGAGTACATGTTCCCTGATGC 540
Qy 541 TGTGAGGTTTCAGGATTCGAGGAGTGAAGGAGTATTCCTTTCTACGAGGATGGCA 600
Db 541 TGTGAGGTTTCAGGATTCGAGGAGTGAAGGAGTATTCCTTTCTACGAGGATGGCA 600
Qy 601 CCAAGAGTGTTCAGGAGTGAAGAGTGAAGGCTTTTAAGGAGGCGCCCTTAAGGCTCTCAG 660
Db 601 CCAAGAGTGTTCAGGAGTGAAGAGTGAAGGCTTTTAAGGAGGCGCCCTTAAGGCTCTCAG 660
Qy 661 AGCATGATTAACCTGGTCAGAGGAAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGT 720
Db 661 AGCATGATTAACCTGGTCAGAGGAAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGT 720
Qy 721 TGTTCAGGTTGATCCCGCTTTTGAAGCAATGGAATGGATGGAAATGGAAAGCTTGGTGAAGTG 780
Db 721 TGTTCAGGTTGATCCCGCTTTTGAAGCAATGGAATGGATGGAAATGGAAAGCTTGGTGAAGTG 780
Qy 781 GAACCCCTGTTCGAAATGTGAAGGCGCATGATGAACTTCTCTTAGGACCATGAATGT 840
Db 781 GAACCCCTGTTCGAAATGTGAAGGCGCATGATGAACTTCTCTTAGGACCATGAATGT 840
Qy 841 GCCTGTGAATTCCTTCGATCAAAAGGATATGTTCCATTTGGGTGTGAGCGCTGCACATAG 900
Db 841 GCCTGTGAATTCCTTCGATCAAAAGGATATGTTCCATTTGGGTGTGAGCGCTGCACATAG 900
Qy 901 GCCTGTTCCTTACCTGGGCAACATGAAGGGAAGGAGGTGGTGGAGGATGCCAAGC 960
Db 901 GCCTGTTCCTTACCTGGGCAACATGAAGGGAAGGAGGTGGTGGAGGATGCCAAGC 960
Qy 961 TAAGGAATGTGGTCTTCACAAAGGAATGTAAAGCAGCAAGAGAGGAGGATGTTAATGG 1020
Db 961 TAAGGAATGTGGTCTTCACAAAGGAATGTAAAGCAGCAAGAGAGGAGGATGTTAATGG 1020
Qy 1021 AAATGGGCTATCCCAATCCCAATGCTGATGCTTACCACTGTGCTGACATTTTCAA 1080
Db 1021 AAATGGGCTATCCCAATCCCAATGCTGATGCTTACCACTGTGCTGACATTTTCAA 1080
Qy 1081 CAGCCGGAATGTAGTTAACTTGAAGGAGGAGTGGAAATGGAATTTGGCAAAATTTGGAGGA 1140
Db 1081 CAGCCCGGAATGTAGTTAACTTGAAGGAGGAGTGGAAATGGAATTTGGCAAAATTTGGAGGA 1140
Qy 1141 CCGAAAGGAACCATGGCTTGTGCTTTATGCAACATGCTGCTTACCTGACGAGCTAT 1200
Db 1141 CCGAAAGGAACCATGGCTTGTGCTTTATGCAACATGCTGCTTACCTGACGAGCTAT 1200
Qy 1201 GGAGGAATCTTATGTGACTTAGCAGCAAGATTTAGCAGGCTCAACAGGAGTGAAGTTGG 1260
Db 1201 GGAGGAATCTTATGTGACTTAGCAGCAAGATTTAGCAGGCTCAACAGGAGTGAAGTTGG 1260
Qy 1261 AAAATTTAGAGCAGATGGAGAAAGAAAGATTTGCAAGAGTGAATGCAATTTGGGAAG 1320
Db 1261 AAAATTTAGAGCAGATGGAGAAAGAAAGATTTGCAAGAGTGAATGCAATTTGGGAAG 1320
Qy 1321 CTTCCCTACCATATATTTTCCCAAGCAATTCGTCTCGGCCCAACATTAAGTATCCCTC 1380
Db 1321 CTTCCCTACCATATATTTTCCCAAGCAATTCGTCTCGGCCCAACATTAAGTATCCCTC 1380
Qy 1381 AGAAAGAGAGATGTTGATTCCTTGTATGGCATTTGTAATGCTTAAAGTATGAGATATCA 1440
Db 1381 AGAAAGAGAGATGTTGATTCCTTGTATGGCATTTGTAATGCTTAAAGTATGAGATATCA 1440
Qy 1441 GGAAATTTCTTCGTTTGGGTGCAATTCATTTGACTATACGTACAGCGGGTTTCCT 1500
```

```
Db 1441 GGAAATTTCTTCGTTTGGGTGCAATTCACCTTTGACTATACGTACAGCGGGTTCT 1500
Qy 1501 TCTTTATGCTATTACGTGATATACCAATTCGTTTACAGATTCTTCTGTGAACCTGTTGGA 1560
Db 1501 TCTTTATGCTATTACGTGATATACCAATTCGTTTACAGATTCTTCTGTGAACCTGTTGGA 1560
Qy 1561 AGTGGGAATCGAGGTTTATACAAATAGATACCTACGTTTGAATGGTTTTAAAAA 1620
Db 1561 AGTGGGAATCGAGGTTTATACAAATAGATACCTACGTTTGAATGGTTTTAAAAA 1620
Qy 1621 AAAAAAAA 1629
Db 1621 AAAAAAAA 1629

RESULT 3
US-10-425-114-7777
; Sequence 7777, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7777
; LENGTH: 1730
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700683191_FLI
US-10-425-114-7777

Query Match 56.3%; Score 916.8; DB 7; Length 1730;
Best Local Similarity 80.7%; Pred. No. 8.7e-264;
Matches 1126; Conservative 0; Mismatches 252; Indels 18; Gaps 4;

Qy 41 GCTTTCACTTCTTCAATTTCCGCACCAACTTCCACCTTCCCATCATCGGAACCCAACTT 100
Db 79 GCTGCTGAGCTCTAGTCTTCTCTCTGCTCGATCTTCTCCCGACGCAAAAGCT 138
Qy 101 CCGCAATTTGGGTCAATTAGGATTTCCGGAGGCGCCCATTTGGAGGCGCGTTAATTTCAAT 160
Db 139 CCTCAATCGGTTGTTTCCGTTTCCGGAGAGGTGCTGTTGCTGTTGTTGTTCAAT 198
Qy 161 TTATCTCAAGAGCGAGGCTTTGTAAGCCGCTTAACCGCAACTTCCACCGCAAGGATTC 220
Db 199 GTAACCTCAACGAGCTCTTGTGTGAGGCCACTCAAGCCGCAACCGCAACGCAATCT 258
Qy 221 ATTGTTCTCTCCGACCAACCACTGCTTCTGCTTCTCTGAGAGGAGGAGAGAT 280
Db 259 ATTGTTCTCTCTGAG--CAACTATGCTTCTCTGAGGTTGAGAGGAGGAGAGAT 315
Qy 281 TTTGAAACAGATAGCCAGTATCTCGCAATGCTTCACTCTTCAAAATCATGGATAGAGCC 340
Db 316 TTTGACCAATAGCGAAGACCTTGAATGATCTCTCTTGAATATGATAGAGGCC 375
Qy 341 CTCGACAAATTCGGCAACGACATAGCTATGCTTCCCTCAGTGGTCTGAAAGATGTTGCTTG 400
Db 376 CTCGACAAATTTGGGAAACGACATCGCTATTTGCAATTTAGTGGTCTGAAAGATGTTGCTTG 435
Qy 401 ATTGAGTATGCGAAATTTGACGGGTCCACCTTTTAGGGTTTTCAGTTTGGACATGGGAGA 460
Db 436 ATTGAGTATGCAATTTGACGGGTCCACCTTACAGAGTGTTTAGTCTTGAACATGGGAGA 495
Qy 461 CTGAACCCAGAACTTATCAACTTTTGTATGCGGTTGAGAGGATTTATGGAATTCGCAATT 520
```

Db 496 CTGATCAGAAACCTACAAATTTTGTGCGGTGAGAAACATATGAAATTCATTT 555
Qy 521 GAGTACATGTTCCCTGATGCTGTTGAGGTTGAGGATGAGGATGAGGATGATTC 580
Db 556 GAGTACATGTTCCCTGATGCTGTTGAGGTTGAGGATGAGGATGAGGATGATTC 615
Qy 581 TCCTTTACGAGGATGGCCACCAAGATGTTGACAGGTTGAGGATGAGGATGAGG 640
Db 616 TCATTTTACGAGGATGGCCATCAAGATGTTGACAGGTTGAGGATGAGGATGAGG 675
Qy 641 AGGCGCTTAAAGGCTCTCAGAGCATGATACTGCTCAGAGGATGAGGATGAGGAT 700
Db 676 AGAGCCCTTAAAGGCTCTCAGAGCATGATACTGCTCAGAGGATGAGGATGAGG 735
Qy 701 ACTAGGCTGTAATACACCGGTTGTTGAGGTTGATCCGCTTTGAGGAAATGATG 760
Db 736 ACTAGGCTGTAATACACCGGTTGTTGAGGTTGATCCGCTTTGAGGAAATGATG 795
Qy 761 ATTGGAAGCTTGGTGAATGGAACCCCTGTTGCAATGTTGAGGATGAGGATGAGG 820
Db 796 ATTGGAAGCTTGGTGAATGGAACCCCTGTTGCAATGTTGAGGATGAGGATGAGG 855
Qy 821 TTCTTTAGGACCATGAAATGCTGCTGTAATTCCTTGCATGCAAAAGGATATGTTCC 880
Db 856 TTCTTTAGGACCATGAAATGCTGCTGTAATTCCTTGCATGCAAAAGGATATGTTCC 915
Qy 881 GGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
Db 916 GGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
Qy 941 TGGTGGAGGATGCAAAAGCTAAGAAATGAGGATGAGGATGAGGATGAGGATGAGG 1000
Db 976 TGGTGGAGGATGCAAAAGCTAAGAAATGAGGATGAGGATGAGGATGAGGATGAGG 1035
Qy 1001 AAAGAGGAGGATGTAATGGAATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1060
Db 1036 GATGCTGCCAGCTTAATGGAATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1086
Qy 1061 ACTGTGCTGACATTTTCAACAGCCCGGATGAGGATGAGGATGAGGATGAGGATGAG 1120
Db 1087 ACTGTGCTGACATTTTCAACAGCCCGGATGAGGATGAGGATGAGGATGAGGATGAG 1146
Qy 1121 AATTGGCAAAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1180
Db 1147 AATTGGCAAAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1206
Qy 1181 TGCCCTACTGCCAGGCTATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1240
Db 1207 TGCCCTACTGCCAGGCTATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1266
Qy 1241 TCAAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1300
Db 1267 T---CAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1323
Qy 1301 AGTGAATGCAATTTGGGAGCTTCCCTAGGATATATTTTCCCAAGCATGCTGCTCGG 1360
Db 1324 ACTGAATGCAATTTGGGAGCTTCCCTAGGATATATTTTCCCAAGCATGCTGCTCGG 1380
Qy 1361 CCAACATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAAT 1420
Db 1381 CAACCAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAAT 1440
Qy 1421 GCCTTAAGATGAGGAT 1436
Db 1441 GCTTTACGGTATGATG 1456

RESULT 4
us-10-424-599-134742
; Sequence 134742, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Zhou Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 134742
; LENGTH: 2774
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92682C.1
US-10-424-599-134742

Query Match 56.3%; Score 916.8; DB 7; Length 2774;
Best Local Similarity 80.7%; Pred. No. 1.2e-263;
Matches 1126; Conservative 0; Mismatches 252; Indels 18; Gaps 4;
Qy 41 GCTTTCATCTTTCATATTTCCGACCAACTTCCACCTTCCCATCATCGGAACCAACTT 100
Db 145 GCTGCTCAGCTTCTAGCTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Qy 101 CCGCAAAATTTGGTCAATTAGGATTTCCGAGAGGCCCATTTGGAGGCGCGTAAATTTCAAT 160
Db 205 CCTCAAAATCGGTTTCGTTTCGGTTTCCGAGAGGTCGCTTGTCTCTGTTGTTGTCAT 264
Qy 161 TTATCTCAAAAGACGGAGCTTGGTAAAGCCGTTTAAACCCGAACTTCCACGCAAGGATTTCC 220
Db 265 GTAACTCAACGACCTCTTGGTGAGGCCACTCAACCCGCAACCCGCAACGATTTCT 324
Qy 221 ATTGTTCTCTCGCAGCAACCAACCATCGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 280
Db 325 ATTGTTCT 381
Qy 281 TTTGAAAGATAGCCAGTGTATCTGCAATGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 340
Db 382 TTTGAGCAAAATAGCGAAGACCTTTGAAATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 441
Qy 341 CTGCAAAAATTTGGCAACGACATAGCTATTTGCTTTCAGTGTGTCTGAAAGATGTTGCTTTG 400
Db 442 CTGCAAAAATTTGGCAACGACATAGCTATTTGCTTTCAGTGTGTCTGAAAGATGTTGCTTTG 501
Qy 401 ATTGAGTATGCAAAATTTGACGGGTGACCCCTTTAGGGTTTTCAGTTTGGGACACTGGGAGA 460
Db 502 ATTGAGTATGCAAAATTTGACGGGTGACCCCTTACAGAGTGTCTCTCTCTCTCTCTCTCTCT 561
Qy 461 CTGAACCCAGAACTTATCACTTTTGTGATGCGGTTCGAGAGCATTTAGGAAATTCGCATT 520
Db 562 CTGAATCCAGAACTTATCACTTTTGTGATGCGGTTCGAGAGCATTTAGGAAATTCGCATT 621
Qy 521 GAGTACATGTTTCCCTGATGCTGTTGAGGTTTTCAGGATTTGGTGGAGGATTAAGGGGTTATTC 580
Db 622 GAGTACATGTTTCCCTGATGCGGTTGAGGTTTTCAGGATTTAGTAACTTAAGGGGCTCTTC 681
Qy 581 TCCTTTCTACGAGGATGGGCACCAAGAGTGTGTCAGGGGTGAGAAAGGTTGAGGCGCTTTAAGG 640
Db 682 TCATTTTACGAGGATGGGCATCAAGAGTGTGTCGAGTGAAGGATGAGGCGCTTTGAGG 741
Qy 641 AGGCGCTTAAAGGCTCTCAGAGCATGATATCTGCTCAGAGGATGAGGATGAGGATGAGGATGAG 700
Db 742 AGAGCCCTTAAAGGCTCTCAAAAGCATGGATCTAGGACAGAGAAAGATTCAGTCTCTCTCT 801
Qy 701 ACTAGGCTGAAATACCGGTTGTTGAGGTTGATCCGCTTTTGGGGGATGAGGATGAGGATGAG 760
Db 802 ACTAGGCTGAAATACCGGTTGTTGAGGTTGATCCGCTTTTGGGGGATGAGGATGAGGATGAG 861
Qy 761 ATTGGAAGCTTGGTGAATGGAACCCCTGTTGCAATGTTGAGGATGAGGATGAGGATGAGGATGAG 820
Db 862 ATTGGAAGCTTGGTGAATGGAACCCCTGTTGCAATGTTGAGGATGAGGATGAGGATGAGGATGAG 921

```

Qy 821 TTCCTTAGGACCATGAATGTGCTGTGAATTCCTTGATGCAATGCAAAAGGATATGTTTCATT 880
Db 922 TTCCTTAGGACCATGAATGTGCTGTGAATTCCTTGATGCAATGCAAAAGGATATGTTTCATT 981
Qy 881 GGGTGTAGCCCTGCACTAGGCTGTGTTTACCTGGGCAACATGAAAGGAGGAGTGG 940
Db 982 GGGTGTAGCCCTGCACTAGGCTGTGTTTACCTGGGCAACATGAAAGGAGGAGTGG 1041
Qy 941 TGGTGGGAGGATGCCAAGCTTAAGCAATGTGCTTCCACAAAGGAAATGTAAGCAGCAG 1000
Db 1042 TGGTGGGAGGATGCCAAGCTTAAGCAATGTGCTTCCACAAAGGAAATGTAAGCAGCAG 1101
Qy 1001 AAGAGGAGGATGTTAATGGAATGGGCTATGCCAATCCCAATGCCAATGGTGTGCTTACC 1060
Db 1102 GATGCTGCCAGCTTAATGGAATGGGCTTCCCAAGCAAAATGGA-----TCTGCC 1152
Qy 1061 ACTGTGCTGACATTTTCAACAGCCGCAATGTAGTTAACTTGAGCAGACCTGGAATTTGAG 1120
Db 1153 ACTGTGCTGACATTTTCAACAGCCGCAATGTAGTTAACTTGAGCAGACCTGGAATTTGAG 1212
Qy 1121 AATTGGCAAAATTTGGAGGACCCGAAAGCAACATGGCTTGTGTGCTTATGCAACCATGG 1180
Db 1213 AATTGGCAAAATTTGGAGACCCGAAAGCAACATGGCTTGTGTGCTTATGCAACCATGG 1272
Qy 1181 TGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTATGAGCAGCAAGTTAGCAGGG 1240
Db 1273 TGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTATGAGCAGCAAGTTAGCAGGG 1332
Qy 1241 TCAACAGGATGAAGTTTGGAAATTTAGAGCAGATGGAGAACAGAAAGATTTGCAAG 1300
Db 1333 T---CAGGAGTGAAGTTTGGAAATTTAGAGCAGATGGAGAACAGAAAGATTTGCAAG 1389
Qy 1301 AGTGAATGCAATTTGGGAAGCTTCCCTACCATATTTATTTTCCCAAGACATTCGTCTCGG 1360
Db 1390 ACTGAATGCAATTTGGGAAGCTTCCCTACCATATTTATTTTCCCAAGACATTCGTCTCGG 1446
Qy 1361 CCACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCATTGTTAAAT 1420
Db 1447 CAACCAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCATTGTTAAAT 1506
Qy 1421 GCCTTAAGATGAGAT 1436
Db 1507 GCTTTACGGTGTGTT 1522

```

RESULT 5

```

US-10-731-525-5
; Sequence 5, Application US/10731525
; Publication No. US2004012140A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10731,525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-10-731-525-5

```

```

Query Match      55.58; Score 904; DB 7; Length 1795;
Best Local Similarity 81.33; Pred. No. 6.2e-260;
Matches 1103; Conservative 0; Mismatches 235; Indels 18; Gaps 4;

```

```

Qy 81 CATCATCGGAACCCAAATCTCCGAAATGGGTCAATTAGGATTTCCGAGAGGCCCAATTG 140
Db 135 CTTTCATCGGACGCTAAAGCTCCGCAAAATGGTTCCTTTGGTTCCTCGGAGAGGCTCAAG 194
Qy 141 GAGGCGCGTAAATTTCAATTTATCTCAAGACGCGAGCTTGGTAAAGCCCGTTAACGCCG 200
Db 195 TTTCTGCTGGTGTGTTAAATTTAACTCAAAAGACGCTCTCCGTTGAGGCCCACTCAATGCCG 254
Qy 201 AACTTCACCAAGAGATTCATTTGTTCTCTCGCAGCAACCAACCATCGTTGCTTCTGCTT 260
Db 255 AACCGCAACGGAATGATTCGTGTTCTCTCGAG---CAACTATGCTGCTCTCGAGG 311
Qy 261 CTGAGACGAAGAGGAGATTTTGAACAGATAGCCAGTATCTGACAATTCGTTTCACTC 320
Db 312 TTGAGAAGGAGAAAGAGATTTTGGAGCAATTTAGCGAAAGACCTTGAAATTTTCATCTCCTC 371
Qy 321 TTCAAAATCATGGATAGAGCCCTCGAANAATTTCCGCAACGACATAGCTTATTCCTTCACTG 380
Db 372 TTGAGATTTATGGATAAGGCCCTCGAANAATTTGGGAACGACATCGCTTATTCGCTTTAGTG 431
Qy 381 GTGCTGAAGATGTTGCTTTGATTGAGTATGCGAAATTTGACGGGTGCGACCTTTTAGGGTTT 440
Db 432 GTGCTGAAGATGTTGCTTTGATTGAGTATGCAATTTGACGGGTGCGACCTTACAGAGTGT 491
Qy 441 TCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTGATGCGGTTGAGA 500
Db 492 TTAGTCTTGACACTGGGAGACTGAACCCAGAAACTTACAAATTTTGTGACGCTGTTGAGA 551
Qy 501 AGCAATTTGCAATTCGCATTTGAGTACATGTTTCCCTGATGCTGTTGAGTTTCAAGCATTTG 560
Db 552 AGCAATTTGCAATTTGAGTACATGTTTCCCTGATGCTGTTGAGTTTCAAGCATTTG 611
Qy 561 TGAGGAGTAAAGGGGTTATTTCTCTTCTACGAGATGGGCAACCAAGAGTGTGTGAGGGTGA 620
Db 612 TAGAAGTAAAGGGGCTTCTCTCAATTTACGAGATGGGCAATCAAGAGTGTGTGAGATAA 671
Qy 621 GAAAGGTGAGGCTTTTAAAGGAGGCGCTTAAAGGCTCTCAGAGATGGATTAAGTGTGAGA 680
Db 672 GAAAGGTGAGGCTTTTAAAGGAGGCGCTTAAAGGCTCTCAGAGATGGATTAAGTGTGAGA 731
Qy 681 GGAAGACCAAGTCACTGTGACTAGGTCTGAAATATCCGTTGTTTCAAGTTGATCCGCGCTT 740
Db 732 GAAAGACCAAGTCACTGTGACTAGGTCTGAAATATCCGTTGTTTCAAGTTGATCCGCGCTT 791
Qy 741 TTAGGAGAAATGGAATGTTGGAATTTGGAAGCTTGTGGAAGTGGCAACCCCTGTTGCAAAATCTGA 800
Db 792 TTGAGGAGCTGGAATGTTGGAATTTGGCAGCTGTGGAAGTGGCAACCCCTGTTGCAAAATCTGA 851
Qy 801 AGGGCCATGACATATGGAATCTCCTTTAGGACCAATGAAATGTCCTGTGAAATTCCTTTGCAATG 860
Db 852 ATGGTCTAGACATATGGAATCTCCTTTAGGACCAATGAAATGTCCTGTGAAATTCCTTTGCAATG 911
Qy 861 CAAAAGGATATGTTTCATTTGGGTGAGCCCTGCACTAGGCTGCTTTTACCTTGGGCAAC 920
Db 912 CCCAAGGATATGTTTTCGATTTGGCTGTGAGCCTATGCAAGGCGCGGTTTATCCCGGCAAC 971
Qy 921 ATGAAAGGGAAGGAGGCTGTGCTGGAGATGCCAAAGCTAAGGAATGTGCTTCTCACA 980
Db 972 ATGAAAGGGAAGGAGGCTGTGCTGGAGATGCCAAAGCTAAGGAATGTGCTTCTCACA 1031
Qy 981 AAGGAAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATTCCTCAATPCC 1040
Db 1032 AAGGTAATTTGAAACAGGAAGATGCTGCCAGCTTAATGGAATGGGACCTCCCAAGGAA 1091
Qy 1041 ATGCAAAATGTTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTTAACT 1100
Db 1092 A-----TGGCTCTGCCACTGTTGCTGACATTTTTCATCTCTCCAGAAATGTGGTCACT 1142
Qy 1101 TGAGCAGGACTGGAAATTTGCAAAATTTGGAGGACCCGAAAGGAAACCATGGCTTG 1160
Db 1143 TGAGCAGGCTCCGGGATTTGGAATTTGGCAAAATTTAGAGAACCCGAAAGAACTATGGCTTG 1202
Qy 1161 TTTGTGCTTTATGCAACCATGTTGTCCTTCTGAGGCTATGAGGAAATCTTATGTTGACT 1220

```

[illegible]

RESULT 6

```

US-10-762-049-5
; Sequence 5, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-5

```

Query Match	55.5%	Score 904;	DB 7;	Length 1795;
Best Local Similarity	81.3%;	Pred. No. 6.2e-260;		
Matches 1103;	Conservative 0;	Mismatches 235;	Indels 18;	Gaps 4;
Qy	81	CATCATCGGAACCCAAACTTCGCGAAATTCGGTCAATTAGGATTTCCGAGAGGCCCATTTG	140	
Db	135	CTTCATCGGACGCTAAAGCTCCGAAATTCGTTCTTCGTTTCGGAGAGGCCCTCAAG	194	
Qy	141	GAGGCGCGTTAAATTTCAAATTTATCTCAAAGACGGAGCTTCGTAAGAGCCCGTTAAACGCG	200	
Db	195	TTTCGTCGTGGTGTGTTAAATTTAACTCAAGACGCTCTCGTGAGGCCACTCAATGCCG	254	
Qy	201	AACCTCCACGACGAGGATTCCAATTGTTCTCTCGCAGCAACAACCATCGTTCTGCTTCTGCTT	260	
Db	255	AACCGCAACGGAATGATTCGTGTTCTCTTTGCAG ---CAACTATCGTTGCTCCTCGAGG	311	
Qy	261	CTGAGACGAAAGAGGAAGATTTTGAAACAGATAGCCAGTGTCTCGACAAATGCTTCACCTC	320	
Db	312	TTGAGAGGAGAAAGAGATTTTGAGCAATTTAGCGAAAGACCTTGAAATTCATCTCCTC	371	
Qy	321	TTGAATCATGGATAGAGCCCTCGACAAATTCGGCAGCATAGCTATTGCGCTTCAGTG	380	
Db	372	TTGAGATTATGGATAAGCCCTCGAGAAATTTGGGAACGACATCGCTATTGCCCTTTAGTG	431	
Qy	381	GTGCTGAAGATGTTGCTTTGATTGAGTATCGAAATTCACGGGTTCGACCCCTTAGGGGTTT	440	
Db	432	GTGCTGAAGATGTTGCTTTGATTGAGTATGCACATTTGACGGGTTCGACCCCTACAGAGTGT	491	
Qy	441	TCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTTTGATCGGTTGAGA	500	

Db	492	TTAGTCTTGACACTGGGAGACTGAACCCAGAAACCTTCAAAATTTTTTTTGACGCTGTTGACA	551
Qy	501	AGCATTTATGAAATTCGCAATTGAGTACATGTTCCCTGATGCTGTGTGAGGTTTCAGGCATTTGG	560
Db	552	AGCATTTATGAAATTCATATTTGAGTACATGTTCCCTGATGCGTTGAGGTTTCAGGCATTTAG	611
Qy	561	TGAGGAGTAAAGGGTTATTCTCTTTCTACGAGATGGGCACCAAGAGTGTGTGACGGGTGA	620
Db	612	TAAAGAACTAAGGGGCTCTTCTCATTTTTACGAGGATGGGCATCAAGAGTGTGTGAGAGTAA	671
Qy	621	GAAAGTGTAGGCGCTTTAAAGGAGGCGCTTAAAGGCTCTCAGAGCATGGATAACTTGGTCAGA	680
Db	672	GAAAGTGTAGGCGCTTTGAGGAGAGCCCTTAAAGGCTCTCAAGCATGGATTAATTCAGGACAGA	731
Qy	681	GGAAAGACCAAGTCACCTGGTACTAGGTCTGAAATACCGGTGTGTTCAGGTTGATCCGGCTT	740
Db	732	GAAAGACCAAGTCCTCTGGTACTAGGTCTGAAATCCCTATTGTCCAGGTTGATCCTGTTT	791
Qy	741	TTGAGGGAATGGATGGTGGAAATTTGGAAGCTTGGTGAAGTGGAACCCCTGTTTGCAAATGTGA	800
Db	792	TTGAGGACCTGGATGGTGGAAATTTGGCAGCCTGGTGAAGTGGAACCCCGGTTTGCAAATGTTA	851
Qy	801	AGGGCATGACATATGSAACCTCTCTTAGGACCATGAATGTGCCGTGTGCAATTCCTTGTGATG	860
Db	852	ATGGCTTAGACATATGNACTTCTTAGGACCATGAATGTTCCTGTAAATTCATTGCAATT	911
Qy	861	CAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCCCTGTTTTTACCTGGGCAAC	920
Db	912	CCCAAGGATATGTTTGGATTTGGCTGTGAGCATGCACAAGGCCGGTTTTTACCCGGACAAC	971
Qy	921	ATGAAAGGGAAGGAGTGTGTGTGGGAGATGCCAAAGCTTAAGGATGTGGTCTTCACA	980
Db	972	ATGAAAGAGAAGGAGTGTGTGTGGGAGATGCCAAAGCCAAGAGTGTGGTCTTCACA	1031
Qy	981	AAGGAATGTAAAGCAGCAGAAAGAGGAGATGTTAATGGAATGGGCTATCCCAATCCC	1040
Db	1032	AAGGTAAATTTGAAACAGGAAGATGCTGCCAGCTTAATGGAAATGGGACCTCCCAGGAA	1091
Qy	1041	ATGCAAAATGGTGATGTCAACATGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACT	1100
Db	1092	A-----TGGCTCTGCCACTGTTGCTGCACATTTTCACTCCCAAGATGTGGTTCAGCT	1142
Qy	1101	TGAGCAGGACTGGAAATTTAGAAATTTGGCAAAATTTGGAGGACCGGAAGGACCATGGCTTG	1160
Db	1143	TGAGCAGGCTCGGGATTTGAGAAATTTGGCAAAATTTAGAGAACCCGAAAAGAACACCTGGCTTG	1202
Qy	1161	TTGTGCTTTATGCACCACTGTGCCCTTACTGCGCAGGCTATGGAGGAATCTTTATGTTGACT	1220
Db	1203	TTGTGCTCTATGCACCATGTGTCCGCTTCTGTGAGGCTATGGAGGAGTGGTATGTTGATC	1262
Qy	1221	TAGCAGACAAGTTAGCAGGGTCAACAGGGAAGAAAGTTGGAAAAATTTAGAGCAGATGGAG	1280
Db	1263	TGGCAGAGAAGTTAGCAAGGT---CAGGAGTGAAGTTGCAAAATTCAGAGCCGATGGAG	1319
Qy	1281	AA CHAAGAAGAAATTTGCAAGAGTGCAACTGCNAATTTGGGAGCTTCCCTACGATATATTTT	1340
Db	1320	AGCAGAAGGAATATGCAAGAGTGAACTGCAGTTTGGGAAGCTTCCCCACAATCTTCTCT	1379
Qy	1341	TCCCAAGACATTTGCTCTCGGCCAACAAATAAGATATCCCTCAGAAAAAGAGAGATGTTGATT	1400
Db	1380	TCCCAAGACATCTTCT---CAACCAATTAAGTACCCTTCAGAAAAAGAGAGATGTTGATT	1430
Qy	1401	CTTTGATGGCAATTTGTAATGCCCTTAAGATGAGGAT	1436
Db	1437	CATTGACGGCAATTCGTGAATGCCCTTACGGTGAATG	1472

RESULT 7
US-10-424-599-135400
; Sequence 135400, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

Db 667 ACTGGTCAGAGGAAGATCAATCTCCGGGCAAGGCTCAGATTCGGTGTTCAGGTT 726
Qy 731 GATCCGGCTTTTGAAGGAAATGGATGGTGGAAATTTGGAAAGCTTGGTGAAGTGGAAACCCCTGTT 790
Db 727 GATCCGGCTTTGAAGGTTTGGATGGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 786
Qy 791 GCAAAATGTAAGGCCATGACATATGGAATCTCTTAGGACCATGAATGTGCCTGTGAAT 850
Db 787 GCGAATCTTGAAGGAAATGATGTTTGGAACTCTTGAAGGACTATGGATGTTCCGGTTAAC 846
Qy 851 TCCTTGATCATCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGACCTAGGCTGTTTTA 910
Db 847 ACATTGATCGGCAGGGTATATATCGATTGGATGTGAGCCCTGCACAAAGCGTTTTA 906
Qy 911 CCTGGGCAACATGAAGGGAGGGAGGTGGTGGGAGGATGCAAAAGCTTAAGGAATGT 970
Db 907 CCGGTCAGCACGAGAGAGGAGATGGTGGTGGGAAGATGCTAAAGCCAAGGAATGT 966
Qy 971 GGTCTTCAAAAGAAATGTAAAGCAGACAGAAAGAGGAGATGTTAATGGAATGGGCTA 1030
Db 967 GGACTTCACAAAGGAAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1023
Qy 1031 TCCCAATCCCATGCAAAATGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAAT 1090
Db 1024 TCGAAAT-----CCGCTGTGCAGATATCTTTAAGATGAGAT 1062
Qy 1091 GTAGTTAACTTTAGCAGCACTGGAATTTGAGAAATTTGGCAAAATTTGGAGACCCGAAGAA 1150
Db 1063 CTTGTGACTTTGACGAGCAGGGGATTTGAGAAATTTGATGAATTTGGAGAACCGTAAAGAG 1122
Qy 1151 CCATGGCTTGTGTGCTTTATGCAACATGTGTGCCCCCTACTGCCAGGCTATGGAGGAATCT 1210
Db 1123 CTTTGGATCGTGTGCTTTATGCTCCGCTGTGTGCCCCCTTTTGTCAAGCCATGGAAGCATCG 1182
Qy 1211 TATGTTGACTTAGCAGCAAGTTAGCAGGTCACAGGATGAAGTTTGGAAATTTTGA 1270
Db 1183 TATGATGAACATGGCGGATTAATTTGGCTGG---AAGTGGATTAAGTTTGCACAAATTCAGA 1239
Qy 1271 GCAGATGAGAACAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACG 1330
Db 1240 GCAGATGTTGACCAAGAGGATTTGCTTAAGCAGGAATTCAGCTCGGTAGCTTCCCTACC 1299
Qy 1331 ATATTATTTTCCAAAGCATTGCTTCGGCCCAACATAAAGTATCCCTCAGAAAGAGA 1390
Db 1300 ATTCTGGTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGCTCTGAGAAGAGA 1356
Qy 1391 GATGTTGATTCCTTGATGGCATTTGTAATGCCTTAAGATGA 1432
Db 1357 GATGTTGAGTCTTTGACTTCTTGAATCTTTGTCGGATAA 1398

RESULT 9
US-09-938-842A-498
; Sequence 498, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 498
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-498
Query Match 40.0%; Score 651.6; DB 3; Length 1398;
Best Local Similarity 70.9%; Pred. No. 3.6e-184; Indels 42; Gaps 6;
Matches 966; Conservative 0; Mismatches 354;
Qy 83 TCATCGGAACCCAAATCTCCGAAATTTGGTCAATTAGGATTTTCGGAGAGCCCAATTGGA 142
Db 67 TCATCGAGCCAAAGTTTCGAAATTTGGTTCGTGTAGGTTAATTGGATCGTGTTCATGTT 126
Qy 143 GGGCCCGTTAATTTCAATTTATCTCAAGACGG---AGCTTGGTAAGCCCGTTAAAGCC 199
Db 127 GCTCCTGTGTCTCGAATCTATCTGGGAAGCGATCATCTCTGTAAACCTTTAAAGCGCT 186
Qy 200 GAACCTCCACGCAAGGATTCATTTGTTCTCTCGCAGCAACAAACATCTGTTCTGCT 259
Db 187 GAAACCAAGACAAGGATTCATGATTCCTCTTGGCGCAACATGGTAGCAGAAATTTGCA 246
Qy 260 TCTGAGACGAAAGAG-----GAAGATTTTGAACAGATAGCAGTGTCTCGACAAT 310
Db 247 GAGGAAGTTCAAGTGGTTGAGATTTGAGATTTTGAAGAGCTTGTCTAAGAAAGTTAGAGA 306
Qy 311 GCTTCACCTCTTGAATCATGATAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
Db 307 GCTTCACCTCTGAGATTTATGGCAAAAGCTCTTGAGAAATACGGGAACGATATCGCCAT 366
Qy 371 GCCTTCAGTGTGCTGAAGATTTGCTTTGATTTAGTATGCGAAATTTGAOGGTTCGACCC 430
Db 367 GCATTTAGTGTGCAAGAGATTTGCTCTTTATTGAGTACGCTCATTTGACATGGGAGGCCA 426
Qy 431 TTTAGGGTTTTCAAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTAT 490
Db 427 TTTAGAGTATTTAGTTTGGATACAGGGAGGTTGAATCCTGAGACGATATCGGTTTTTCGAT 486
Qy 491 CGGTTGAGAAGCATTTATGGAATTCGCATTTGAGTACATGTTCCCTGATGCTGTGTGAGGTT 550
Db 487 CGGTTGAGAAGCACTATGGGATTTAGGATTTAGTATATGTTTCTGATTTCTGTTGAGGTT 546
Qy 551 CAGCATTTGTTGAGGAGTAAAGGGTATTTCTTTTCTACAGAGATGGGCAACCAAGAGTGT 610
Db 547 CAAGTTTGGTTAGACACAGGATTTGTTCTCTTTTATGAGGATGGTTCATCAGAGTGT 606
Qy 611 TGCAAGGTTGAGAAAGGTGAGGCCCTTTAAAGAGGGCCCTTTAAGGGTCTCAGAGCATGGATA 670
Db 607 TGCCGTGTTGAAAGGTGAGACCTTTGAGGGCTGCTCTCAAGGGTTTAAAGGCTTGGATT 666
Qy 671 ACTGTCAGAGGAAGAACCACTGACCTGGTACTAGGTTCTGAAATACCGGTTGTTTCAGGTT 730
Db 667 ACTGTCAGAGGAAGAAATCAATCTCCGGGCAAGAGTCTGAGATTTCCGGTTGTTTCAGGTT 726
Qy 731 GATCCGGCTTTTGAAGGAAATGGAATTTGGAAGCTTTGGAAGTGGTGAAGTGGAAACCCCTGTT 790
Db 727 GATCCGGTGTTTGAAGGTTTGAATTTGGATTTGGTGTGGTGTGGTGAAGTGGAAATCCGGTT 786
Qy 791 GCAAAATGTGAAGGCCCATGACATATGGAATCTCTTAGGACCATGAATGTGCCTGTGTGAAT 850
Db 787 GCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTGAAGGACTATGAGATGTTTCCGGTTAAC 846
Qy 851 TCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGACCTAGGCTGTTTTA 910
Db 847 ACATTGATCGGCAGGGTATATATCGAATTTGAGCTGTGAGCCCTTGACCAAGACGGTTTTA 906
Qy 911 CCTGGCAACATGAAGGGAGGAGGTGGTGGGAGAGGATGCGCAAAAGCTTAAGGAATGT 970
Db 907 CCGGTCAGCACGAGAGAGGAGATGGTGGTGGGAAGATGCTTAAGCCAGGAATGT 966
Qy 971 GGTCTTCAAAAGAAATGTAAAGCAGACAGAAAGAGGAGATGTTAATGGAATGGGCTA 1030
Db 967 GGACTTCACAAAGGAAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1023

Db 300 CATGGATAAGAGCTTTGAGAGATTGAGAGACCAATCGCAATTCCTTTAGTGAGAGCTGA 359
QY 388 AGATGTTGCTTTGATGAGATATGCGAAATTTGACGGGTCGACCCCTTTAGGGTTTTTCAGTTTT
Db 360 AGATGTTGCAATGATTAATGATGACCTTTAACTGGAAGCCATTTAGGGTTTTTAGTTTT 419
QY 448 GGACATCTGGGAGACTGAACCCAGAACTTATCAACTTTTGTGATGCGGTTGAGAGCAATTA 507
Db 420 AGATACAGGAGATTAACCCCTGAACGTCAGGCTCTTTGACGAGTCGAGAGACGATGA 479
QY 508 TGGAAATTCGCAATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAAGGCTTTGAGGAG 567
Db 480 CGGATTCGAATTCAGTACATGTTCTCTGATGAGTTGAGGTTTCAAGGCTTTAGTGAGGAA 539
QY 568 TAAGGGGTTATTCCTTTTACGAGGATGGGCAACAGAGTGTTCAGGGTGAGAAAGGT 627
Db 540 CAAGGGTTTCTCTCTATGAGATGCTCAATCAAGAGTGTTCGAGGTTGAGGAAAGT 599
QY 628 GAGGCTTTTAAGGAGGCTTAAAGGCTCTCAAGGCTCTCAAGGCTTCAAGGCTTCAAGGCTT 687
Db 600 TAGACCTTTGCGTCTCTTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGGCT 659
QY 688 CCAGTCACTGCTACTAGGCTCTGAATACCGGTTGTTTCAAGGCTTCAAGGCTTCAAGGCT 747
Db 660 CCAATCTCGGTTACGAGATCTGATCCCTATTTGTTTCAAGGCTTCAAGGCTTCAAGGCT 719
QY 748 AATGATGTTGGAATTCGAAGCTTTGAGGCTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 807
Db 720 GTTAGATGGGCTGTTGGAAGCTCTGTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGGCT 779
QY 808 TGACATATGGAATCTTCTAGGACCATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGAT 867
Db 780 TGATGTTGGAATCTTCTGGAAGCTCTGTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGGCT 839
QY 868 ATATGTTTCCATTCGGTGTGAGGCTCTGATGAGGCTTGTGATGAGGCTTGTGATGAGGCT 927
Db 840 GTATGTTCAATCGGTTGAGGCTGCTGATGAGGCTTGTGATGAGGCTTGTGATGAGGCT 899
QY 928 GGAAGGAGGTGTTGTTGGGAGGATGCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 987
Db 900 AGAAGGAGGTGTTGTTGGGAGGATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 959
QY 988 TGTAAAGCAGAGAAAGGAGGATGTTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 1047
Db 960 CATCAAGAGGAGATGCTG-----CTGAGA 986
QY 1048 TGGTGTGCTACCACTGCTGCTGACATTTTCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
Db 987 CTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
QY 1108 GACTGGAATTTGAGAAATTTGGAAGGATGGAAGGATGGAAGGATGGAAGGATGGAAGGATGGA 1167
Db 1047 AGGAGGGGTTGAGAATCTTTTGAAGCTTAGAAGGATGGAAGGATGGAAGGATGGAAGGATGGA 1106
QY 1168 TTATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Db 1107 TTACGCTCTTGGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
QY 1228 CAAGTTAGCAGGTTCAACAGGATGAGGTTTGAAGGATGAGGATGAGGATGAGGATGAGGATGAG 1287
Db 1167 GAACTTGGGG-----AAGAGGATTAAGTGGGAGATTAAGTGGGAGATTAAGTGGGAGATTAAG 1223
QY 1288 AGAATTTGCAAGAGTGAATGCAATTTGGGAGGCTTCCCTGAGATTAATTTTTCCTCCAAA 1347
Db 1224 GGAGTTTCTTAAGCAAGAGCTTCAAGTTAGGAGGCTTCCCTGAGATTAATTTTTCCTCCAAA 1283
QY 1348 GCATTCGCTCGGCCCAACAATTAAGATATCCCTCAGAAAAGGAGATGTTGATTCCTTGAT 1407
Db 1284 AAGAGCTTCAAGG-----GCTATTAAGTATCCCTCAGAGCATGAGATGTTGATTCCTCAT 1407
QY 1408 GGCATTTTGAATGCTTTAAGATGA 1432

Db 1341 GTCGTTTGTGAATCTTCTTCGGTGA 1365

RESULT 12

US-09-938-842A-2305
; Sequence 2305, Application US/09938842A
; Publication No. US20040009476A9

GENERAL INFORMATION:

; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2305
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2305

Query Match 35.4%; Score 577.4; DB 3; Length 1365;
Best Local Similarity 68.9%; Pred. No. 7e-162;
Matches 886; Conservative 0; Mismatches 351; Indels 48; Gaps 5;

QY 157 CAATTATCTCAAGACGAGCTTTGTAAGCCGGTTAAAGCCGTTAAAGCCGAA---ACCTCACGCA 213
Db 120 CCATTTGCTCAGAGAGCTTACTCTATGAAACCTTTAAAGCTGAGTCACATTCACGAAG 179
QY 214 GGATTCATTTGCTCTCGCAGCAACACATCGTTGCTTCTGCTTCTGAGA-----C 267
Db 180 CGAATCTTGGGTTACTCGTCTTCACTCTAATTTGCTTCTGAAAGTTGAGGAAAGGAGG 239
QY 268 GAAAGAGGAGATTTTGAACAGATAGCAGTGTCTCGCAATGCTTCACTCTTGAAT 327
Db 240 AGAATTTGAAGACTTTGAGCAACTTGTCTAAAGCTTGAAGATGCTTCTCCATTTGAAT 299
QY 328 CATGATAGAGCCCTCGACAAAATTCGGAACAGCATAGCTATTGCTTCACTGCTGCTGA 387
Db 300 CATGGATAAAGCTCTTGAGAGATTCGAGAGACCAATCGCAATTCCTTTAGTGAGCTGA 359
QY 388 AGATGTTGCTTTGATTCAGTATGCAAAATTCAGCGGTCGACCTTTAGGGTTTTTCAGTTT 447
Db 360 AGATGTTGCTTTGATTTGAATATGACGTTTACTTGAAGAGCATTTAGGGTTTTTAGTTT 419
QY 448 GGACATCTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGCGGTTGAGAGCAATTA 507
Db 420 AGATACAGGAGATTAACCCCTGAAACGTCAGGCTCTTTGACGAGTCGAGAGACGATGA 479
QY 508 TGGAAATTCGCAATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAAGGCTTGGTGAGAG 567
Db 480 CGGATTCGAATTCAGTACATGTTTCTCTGATGCTGAGTTTGAAGTTTCAAGCTTTAGTGAGGAA 539
QY 568 TAAGGGGTTATTCCTTTTACAGGAGTGGGCAACAGAGTGTTCAGGGTTGAGAGAGGT 627
Db 540 CAAGGGTTTGTCTCATTTATGAAGATGCTTCAAGAGTGTTCAGAGTGTTCAGGAAAGT 599
QY 628 GAGGCTTTTAAGGAGGCGCTTAAAGGCTCTCAGAGCATGGAATACTGCTCAGAGAGAAAGA 687
Db 600 TAGACCTTTGCGTCTGCTCTTTAAGGCTCTTAAAGCTTGAAGTTTGAAGTACAGGACAGAGGAAAGA 659
QY 688 CCAGTCACTGCTAGGCTCTGAAATACCGGTTGTTTCAAGGTTGATCCGGCTTTTGAAGG 747
Db 660 CCAATCTCGGCTACGAGATCTGATCCCTTATTTGTTTCAAGGTTGATCCAGTGTGTTGAAGG 719

Db 1328 GCGAGCAGAAGCCATTCGCGAGCGAGCTGCAACTCAGAGCTTCCGACGATCCTC 1387
Qy 1337 TTTTCCCAAAGCATTCGTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGAGATGTT 1396
Db 1388 CTGTTCCCGCGCGCACCGGTGAAGC---CCATCAAGTACCGTCCGAGAAGAGGACGTC 1444
Qy 1397 GATTCCTTGATGGCATTTGTAATGCTTAAAGATGAG 1433
Db 1445 CAGTCCCTCTCGCTTCGTGAACAGCCTCAGATGAG 1481

RESULT 14
US-10-762-049-9
; Sequence 9, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-762-049-9

Query Match 33.0%; Score 538.2; DB 7; Length 1827;
Best Local Similarity 68.1%; Pred. No. 5e-150;
Matches 801; Conservative 0; Mismatches 358; Indels 18; Gaps 3;

Qy 257 GCTTCTGAGACGAAGAGGAGATTTTGAACAGATAGCCAGTGTCTCGAATGCTTCA 316
Db 323 GCATCTCGGTGGCGCGTGGACTACGAGGCCCTGGCGAGGAGCTGGTGGGCGCTCG 382
Qy 317 CCTCTGAATCATGATAGACCCCTCGAATAATTCGCAACGACATAGTATGCTTC 376
Db 383 CCGCTGGAGATCATGATCGTGCCTCGACATGTTCCGCTCCGAAATCGCCATCGCCTTC 442
Qy 377 AGTGTGCTGAAGATGTTGCTTTGATTGAGTATCGAAATGACGGTTCGACCTTTAGG 436
Db 443 AGTGTGCGGAGGACGTGGCCCTCATCGAATACGCGAAATGATGAGCGCCCTTCAGG 502
Qy 437 GTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTT 496
Db 503 GTGTTACGCTTGACACTGGGAGTGAACCCAGACATACGAATCTTCGACAAAGTG 562
Qy 497 GAGAAGCATTAAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATG 556
Db 563 GAGAAGCACTATGTTATCCATCGATGATGATGATGATGATGATGATGATGATGATG 622
Qy 557 TTGTTGAGGAGTAAAGGGGTATTTCTTTCTACGAGGATGGCCACCAAGAGTGTTCAGG 616
Db 623 CTTGTGAGGACAAAGGGGCTCTTCTTTCTACGAGGACGACACCAAGAGTGTTCAGG 682
Qy 617 GTGAGAAAGGTGAGGCTTTAAGGAGGCGCCCTTAAGGGTCTCAGAGCATGATGATG 676
Db 683 GTGAGAAAGGTTCGCGCCCTTGAGGAGGCGCCCTCAAGGGCTCAAGGCTGATCAACGG 742
Qy 677 CAGAGGAAGACCACTACCTGGTACTAGGTCTGAAATACCGTGTGTTCAAGTTGATCCG 736
Db 743 CAGAGGAAGGATCAGTCCCTGGCCACGAGCCAGCATCCCTGTTGTTCAAGTTGATCCG 802
Qy 737 GCTTTTGAAGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 796

Db 803 TCTTTTGAAGGGCTGATGCTGAGCGGTAGCTTGATCAAGTGAACCCCTGTGGCTAAT 862
Qy 797 GTGAAGGGCCATGACATATGGAATCTTCTTAGAACCATGAATGTGCTGTGAATTCCTTG 856
Db 863 GTGGATGGCAAGGATATCTGGACCTTCTCAGGACCATGGATGTCCCTGTGAACACCTG 922
Qy 857 CATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACTAGGCTGTGTTTACCTGGG 916
Db 923 CATGCTCAAGGCTACGTCTCCATTGGGTGAGCCGTGCAACAGGCCCGTGTGTCGGGG 982
Qy 917 CAACATGAAGGGAAGGAGGTGTGTGGGAGGATGCCAAAGCTAAAGGAATGTGCTTT 976
Db 983 CAGCAGGAGGGAAGGAGGTGTGTGGGAGGAGCCACGCGCAAGGAGTGGCGCTG 1042
Qy 977 CACAAGGAATGTAAAGCAGCAGAAAGAGAGGAGATGTTAATGGAATGGCTATCCAA 1036
Db 1043 CACAACGGTA-----ACATCGACAAGGAAGGTGAGGCACCCAAAGGTGCGGCTC 1090
Qy 1037 TCCCATGCAAAATGATGCTACCACTGTGCTGACATTTTCAACAGCCCGGAATGTAGTT 1096
Db 1091 AACGGCAACGGCTCGGCCGAGGCGAGGCCGCCAGACATCTTCAGAGCCAGGCCATCGTC 1150
Qy 1097 AACTTGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGAGGAGCCGAAAGAACCTATGG 1156
Db 1151 AACTTCAACCGTCCCGGATCGAGAACCTCTCTGCGGCTCGAGAACCCGCGCGAGCCGTG 1210
Qy 1157 CTGTTGTGCTTTATGACCATGTGTCCTTACTGCGAGCTATGCGAGGATCTTATGTT 1216
Db 1211 CTCACGCTCTCTACGCTCCCTGTCCTTACTGCGAGCAATGAGAGGCGTCTCATGTT 1270
Qy 1217 GACTTAGCAGCAAGTTAGCAGGCTCAACAGGATGAAGTTTGAATAATTTAGAGCAGAT 1276
Db 1271 GAGCTGGCCGAGAAGCTGAGCGCT---GAGGCATCAAGGTGGCCAGTTCCGCGCGGAC 1327
Qy 1277 GGAACACAGAAAGAAATTTGCAAGAGTGAATCTCAATTTGGAAAGCTTCCCTTACGATTTA 1336
Db 1328 GGCAGCAGAAAGCCATTCGCGCAGGCGAGCTGCAACTACAGAGCTTCCCGACGATCCTC 1387
Qy 1337 TTTTCCCAAGCATTCGTCTCGGCCAACTAAGATATCCCTCAGAAAGAGAGATGTT 1396
Db 1388 CTGTTCCCGCGCGCACCGTGAAGC---CCATCAAGTACCGTCCGAGAGAGGAGCGTC 1444
Qy 1397 GATTCCTTGATGGCATTTGTAATGCTTAAAGTGAAG 1433
Db 1445 CAGTCCCTCTCGCTTCGTGAACAGCCTCAGATGAG 1481

RESULT 15
US-10-739-930-2784
; Sequence 2784, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 2784
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1840)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER4907_2
US-10-739-930-2784

Query Match 31.5%; Score 512.6; DB 8; Length 1840;
Best Local Similarity 64.9%; Pred. No. 2.5e-142;

Matches 774; Conservative 0; Mismatches 415; Indels 3; Gaps 1;	
Qy	243 CCATCGTTGCTTCTGCTTCTGAGACGAAAGAGGAAAGATTGTAACAGATAGCAGTCATC 302
Db	303 CCNCGCGCGCGCGTGGCGCGACGAGCGCGGTGACTACGAGGCCCTGGCGCGGAGC 362
Qy	303 TCAGCAATGCTTCACTCTTTGAATCATGATGAGCCCTCGACAAATTCGGCAACGACA 362
Db	363 TTGAGGCGCGCTGCCCTGGAGATCATGATCGGCGCTGGCCATGTTCCGGTCCGAAA 422
Qy	363 TAGCTATTGCTTTCAGTGGTCTGAGATGCTTCTTTGATTGATGATGCGAAATTCACGG 422
Db	423 TCGCCATTCGCTTTCAGCGCGCGGAGGACCTGCGCTGATCGAGTACGCGANGCTGACGG 482
Qy	423 GTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAAC 482
Db	483 GCGCTCCCTTCGGGTGTTTCAGCTTGGACACGCGGGCGGCTCAACCCGGAGACGTACGAGC 542
Qy	483 TTTTGTGATGCGGTGAGAACATATGGAATTCGCATTTGAGTACATGTTCCCTGATGCTG 542
Db	543 TCTTCGACAGGGTGGAGACGACTACGGCATTCGCCATCGAGTACATGTTCCCGGACGCGA 602
Qy	543 TTCAGGTTTCAGGCAATTCGTCAGGAGTAAGGGTTATCTCTTTCTACGAGGATGGGCACC 602
Db	603 GCGAGGTGCAGGAGCTGGTCCGACCAAGGGGCTCTTCTCTTCTACGAGGACGGCCACC 662
Qy	603 AAGAGTGTTCAGGGTGGAGAAAGTGTAGGCTTTTAAGAGGGGCGCTTAAGGGTCTCAGAG 662
Db	663 AGAGTGTCTCCGCTGGCGAAGGTGGCGCCCTTCGCGAGGGCGCTCAGGGGCTCAGGG 722
Qy	663 CATGGAATACTGTGTGAGAGAAAGACAGTCACTCTGTTACTAGGTCTGAAATACCGGTTG 722
Db	723 CGTGGATCACCGGCCAGAGAGAGACCACTCCCGCGCACCAAGGGCCAGCATCCCCCTGG 782
Qy	723 TTCAGGTTGATCGGCTTTTGAGGGAAATGCGATGGTGGAAATGGAAGCTTGGTGAAGTGA 782
Db	783 TCCAGGTGCAGCCCTTCTCTGAAAGGCTTGACGCGGGGGCTGGTAGCTGGTCAAGTGA 842
Qy	783 ACCCTGTGCAAAATGTGAAGGGCGATGACATATGGAATTCCTTAGGACCATGAATGTGC 842
Db	843 ACCCGTGGCCACGTCGACGGCAAGGACATCTGGACTTCTCCGGACCATGGAGCTCC 902
Qy	843 CTGTGAATTCCTTTCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGC 902
Db	903 CTGTCAAGCCCTGCACGCCCGCAGGGCTACGTGTCCATCGGGTCCGAGCGGTGCACCGGC 962
Qy	903 CTGTTTTCCTTGGGCAACATGAAGAGGAGGAGGTGGTGGTGGAGGATGCCAAAGCTA 962
Db	963 CTGTCTCGCGGGGCGACGACGCGGGAAGGGCGGGTGGTGGTGGAGGACGCCAAGGCCA 1022
Qy	963 AGGAATGTGCTCTTCAAAAGGAAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGAA 1022
Db	1023 AGAGTGGCGCTTCCACAAAGGGCAACATCGAAGAGACGGCCAGGCGGCGCGCGCGCGG 1082
Qy	1023 ATGGGCTATCCCAATCCCATGCAAAATGGTATGCTACCACTGTGCTGACATTTTCAACA 1082
Db	1083 CGCGCGCCCGCCAGGACTGCGCAACGGCAACGGCTCGCGCGCGCGCCCGGACATCTTCGAGA 1142
Qy	1083 GCCCGAATGTAGTTAATTGAGCAGGACTGGAAATGAGAAATTTGGCAAAATTTGGAGGACC 1142
Db	1143 GCCCGCGCGCTGCTGCTCCTCACCCGCGCGGGTTCGAGAACCTGTGCGCTGGAGAGCC 1202
Qy	1143 GAAAGGAACCATGGCTTGTGCTTTTATGCACCATGGTGCCCTACTGCGAGGCTATG 1202
Db	1203 GCGCCGAGCGGTGCTCGTGGTCTGTACGCGCCCTGGTGCCGTTCTGCCAGGCCATGG 1262
Qy	1203 AGGAATCTTATGTTGATTTAGCAGACAAGTTAGCAGGGTCAACAGGGATGAAGGTTGGAA 1262
Db	1263 AGGCTCTCTACGTGGAGCTGGCGAGAGGCTGGCGGGGTCGCGCGCGTGAAGGTGCGCC 1322
Qy	1263 AATTTAGAGCAGATGAGAACAGAAAGAAATTTGCAAGAGTGAATCTGCAATTTGGGAGCT 1322
Db	1323 GGTTCGCGCGCGACGCGGAGCAGAAAGCCGTTTCGCGCAGGCGGAGCTGCGAGCTG 1382

Search completed: February 21, 2006, 00:17:31
Job time : 1414 secs

Qy	1323 TCCTACGATATTATTATTTTCCCAAGCATTCGTCTGGCCCAACAATAAAGTATCCCTCAG 1382
Db	1383 TCCCGACCGTCTCTCTGTTCCCGACGCGGACCGCGCGGC---CCATCAAGTACCGTGG 1439
Qy	1383 AAAAGAGAGATGTTGATTCCTTGTATGGCATTTTGTAAATGCTTTAAGATGAGG 1434
Db	1440 AGAAGAGGAGCGTGGACTCGCTCTCGCTTCGTCAACAGCCTCCGCTGAGG 1491

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 21:14:31 ; Search time 540 Seconds
(without alignments)

6403.821 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacgaggagagaccata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.New.*
1: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/prodata/2/pubpna/US05_NEW_PUB.seq.*
7: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq2.*
11: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq3.*
12: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq4.*
13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	568	34.9	1082144	12	US-11-117-187-211
2	41.8	2.6	1161	8	US-10-750-185-59729
3	41.8	2.6	1161	8	US-10-750-623-59729
4	39.8	2.4	4751	7	US-10-511-538-86
5	39.4	2.4	1366	6	US-09-925-065A-685225
6	38.8	2.4	615	6	US-09-925-065A-827921
7	38.8	2.4	622	6	US-09-925-065A-822195
8	38.8	2.4	194553	12	US-11-098-686-8738
9	38.6	2.4	569	6	US-09-925-065A-335502
10	38	2.3	661	6	US-09-925-065A-687471
11	37.6	2.3	535	6	US-09-925-065A-159939
12	37.6	2.3	615	6	US-09-925-065A-827920
13	37.6	2.3	622	6	US-09-925-065A-822196
14	37	2.3	519	6	US-09-925-065A-466709
15	36.8	2.3	459	6	US-09-925-065A-192552
16	36.8	2.3	618	6	US-09-925-065A-423102
17	36.8	2.3	618	6	US-09-925-065A-423103
18	36.8	2.3	618	6	US-09-925-065A-423104
19	36.8	2.3	1690	12	US-11-090-439-43
20	36.6	2.2	600	12	US-11-136-527-6775

21	36.6	2.2	2314	12	US-11-136-527-2679	Sequence 2679, Ap
22	36.6	2.2	10619	8	US-10-240-708-3	Sequence 3, Appli
c 23	36.6	2.2	27902	8	US-10-995-561-13462	Sequence 13462, A
24	36.6	2.2	65931	8	US-10-995-561-13254	Sequence 13254, A
25	36.4	2.2	549	6	US-09-925-065A-819926	Sequence 819926, A
26	36.2	2.2	1503	12	US-11-013-247A-3	Sequence 3, Appli
27	36.2	2.2	1587	12	US-11-013-247A-18	Sequence 18, Appli
28	36.2	2.2	2457	6	US-09-925-065A-545998	Sequence 545998,
29	36.2	2.2	2457	6	US-09-925-065A-545999	Sequence 545999,
30	36	2.2	571	6	US-09-925-065A-556459	Sequence 556459,
31	35.8	2.2	1520	8	US-10-750-185-42946	Sequence 42946, A
32	35.8	2.2	1520	8	US-10-750-623-42946	Sequence 42946, A
c 33	35.6	2.2	200	12	US-11-098-686-816	Sequence 816, App
34	35.6	2.2	554	6	US-09-925-065A-252352	Sequence 252352,
c 35	35.6	2.2	579	6	US-09-925-065A-561598	Sequence 561598,
36	35.6	2.2	100000	12	US-11-124-367A-5027	Sequence 5027, Ap
c 37	35.4	2.2	345	12	US-11-091-883-335	Sequence 355, App
38	35.4	2.2	1478	8	US-10-909-125-1744	Sequence 1744, Ap
c 39	35.2	2.2	622	6	US-09-925-065A-587553	Sequence 587553,
40	35.2	2.2	737	6	US-09-925-065A-60076	Sequence 60076, A
c 41	35.2	2.2	742	6	US-09-925-065A-87642	Sequence 87642, A
42	35.2	2.2	1751	8	US-10-750-185-52091	Sequence 52091, A
43	35.2	2.2	1751	8	US-10-750-623-52091	Sequence 52091, A
c 44	35.2	2.2	3268	8	US-10-793-626-3947	Sequence 3947, Ap
45	35.2	2.2	3275	8	US-10-750-185-38427	Sequence 38427, A

ALIGNMENTS

RESULT 1

US-11-117-187-211
; Sequence 211, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVEN, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:305US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-211

Query Match	34.9%	Score 568;	DB 12;	Length 1082144;
Best Local Similarity	73.5%;	Pred. No. 2e+142;		
Matches	791;	Conservative	0;	Mismatches 255; Indels 30; Gaps 4;
Qy	370	TGCGTTTCAGTGGTGAAGATGTCCTTTGATTGATGTCGAAATTCAGCGGTCCGACC	429	
Db	354529	TGTTTCAGTGGTGAAGATGTCCTTTGATTGATGTCGAAATTCAGCGGTCCGACC	354588	
Qy	430	CTTTAGGTTTTCAGTGGTGAAGATGTCGAAATTCAGCGGTCCGAAATTCATCAACTTTTGA	489	
Db	354589	ATTAGATGTTTTCAGTGGTGAAGATGTCGAAATTCAGCGGTCCGAAATTCATCAACTTTTGA	354648	
Qy	490	TGCGTTTGAAGATGTCGAAATTCAGTGGTGAAGATGTCGAAATTCAGTGGTTCGAGGT	549	
Db	354649	TGCGTTTGAAGATGTCGAAATTCAGTGGTGAAGATGTCGAAATTCAGTGGTTCGAGGT	354708	
Qy	550	TCAGGATTTGGTGAAGATGTCGAAATTCAGTGGTGAAGATGTCGAAATTCAGTGGTTCGAGGT	609	
Db	354709	TCAGGATTTGGTGAAGATGTCGAAATTCAGTGGTGAAGATGTCGAAATTCAGTGGTTCGAGGT	354768	

```
Qy 610 TTCCAGGTTGAGAAAGTGGAGCCCTTTAAGAGGGCCCTTAAGGGTCTCAGAGCATGGAT 669
Db 354769 TTGCCCGTGTTCGAAAGGTGAGACCTTTTGAGCGGTGCTCTCAAGGGTTTAAAGGCTTGGAT 354828
Qy 670 AACTGGTTCAGAGGAAAGACCAAGTCACTCGTGTACTAGGTCTGAAATACCGGTGTTTCAGGT 729
Db 354829 TACTGGTTCAGAGGAAAGATCAATCTCCGGGACAAAGTCTGAGATTCCGGTGTTCAGGT 354888
Qy 730 TGATCCGGCTTTTGAAGGAATGGATGGTGAATTTGAAGCTTGGTGAAGTGGAAACCTGT 789
Db 354889 TGATCCGGTGTTCGAAAGTTCGATGCTGAGTGGAGTTGGTAGTTTGGTGAAGTGGAAATCCGGT 354948
Qy 790 TGCAAAATGTGAAGGCCATGACATATGGAATCTTCTAGGACCATGAATGTCCTGNGAA 849
Db 354949 TGCAGATGTGAAGGGAATGATGTTGGAACTTCTGAGGACTATGGATGTTCCGGTTAA 355008
Qy 850 TTCTTCGATGCAAAAGGATATGTTTTCATTTGGGTGTGAGCCCTGCACCTAGGCTGTTTT 909
Db 355009 CACATTGCATGCGCAGGGTATATATCGATTTGGATGTGAGCTTGCACGAAAGCGTTTT 355068
Qy 910 ACCTGGGCAACATGAAAGGGAAGGAGGTGTGTGGGAGGATGCCAAAGCTAAGGAATG 969
Db 355069 ACCGGTTCAGCAGAGAGAGGAGGATGTGTGGGAGAGTCTAAAGCCAAAGGAATG 355128
Qy 970 TGGTCTTCACAAAGGAATGTAAAGCAGCAGAGAGGAGGATGTTAATGGAATGGCT 1029
Db 355129 TGGACTTCACAAAGGGAATGTCA---AAGAAATCTCCGATGATGCTAAAGTGAACGGGGA 355185
Qy 1030 ATCCCAATCCCATGCAAAATGGTGTGTACCACTGTGTGCTGACATTTTCAACAGCCGGAA 1089
Db 355186 ATCGAAT-----CCGCTGTTCCAGATATCTTTAAGAGTGAGAA 355224
Qy 1090 TGTAGTTAACTTGAGCAGGACTGGAATGGAATTTGGCAAAATTTGAGAGCCGAAAGGA 1149
Db 355225 TCTTGTGACTTTGAGCAGGACGAGGATTTGAGAAATTTGATGAATTTGGAGAACCGTAAAGA 355284
Qy 1150 ACCATGGCTTGTGTGCTTTATGACCAATGTGCTCCCTACTGCCAGGCTATGGAAGGATC 1209
Db 355285 GCCTTGGATGCTGTGCTTTATGCTCGGTGTGCTCCCTTTTGTCAAGCCATGGAAGCATC 355344
Qy 1210 TTATGTTGACTTAGCAGCAAGTTAGCAGGCTCAACAGGGATGAAGTTGGAAATTTAG 1269
Db 355345 GTATGATGAACCTGGCGGATTAATTTGGCTGG---AAGTGGATTAAGTTGGCAAAATTCAG 355401
Qy 1270 AGCAGATGGAGAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGAGACTTCCCTAC 1329
Db 355402 AGCAGATGGTACCAGAGGAGTTTGTCTAAGCAGGAATTCAGCTCGGTAGCTTCCCTAC 355461
Qy 1330 GATATTATTTTCCCAAGCAATTCGTCTCGGCCCAACATAAAGTATCCCTCAGAAAAGAG 1389
Db 355462 CATCTGGTGTTCCTTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGAGAGAG 355518
Qy 1390 AGATGTTGATTCCTTGATGGCAATTTGTAATGCTTTAAGATCAGGATATCAGGAAA 1445
Db 355519 AGATGTTGAGTCTTTGACTTCGTCTTGAACTTCTCCGATGAAGTAAACCCACAAA 355574
```

RESULT 2

```
US-10-750-185-59729/c
; Sequence 59729, Application US/10750185
; Publication No, US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
```

```
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59729
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Bovine 19866881617930
US-10-750-185-59729
```

```
Query Match 2.6%; Score 41.8; DB 8; Length 1161;
Best Local Similarity 56.0%; Pred. No. 0.62;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
Qy 1367 ATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGTGATGGCATTTGTAATGCCCTTA 1426
Db 886 ATAAACATTTCTCACAACTCACACATGTTGATCACCTATTGACATTTATGAAGGAATTTG 827
Qy 1427 AGATGAGGATATCAGGAAATTTTCTCGTTTTGGTGGTTCGAAATTCACATTTGACTATACG 1486
Db 826 TGTGATGAGATTTTGACTATTTCTTATCTCTATAACTTTAAATGGAATCTTGATGAAT 767
Qy 1487 TACAGCGGTTCCCTCTTTTAT 1507
Db 766 TACTGAGGTATTTTTCATTAT 746
```

RESULT 3

```
US-10-750-623-59729/c
; Sequence 59729, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
```

```
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59729
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Bovine 19866881617930
US-10-750-623-59729
```

```
Query Match 2.6%; Score 41.8; DB 8; Length 1161;
Best Local Similarity 56.0%; Pred. No. 0.62;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
Qy 1367 ATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGTGATGGCATTTGTAATGCCCTTA 1426
Db 886 ATAAACATTTCTCACAACTCACACATGTTGATCACCTATTGACATTTATGAAGGAATTTG 827
Qy 1427 AGATGAGGATATCAGGAAATTTTCTCGTTTTGGTGGTTCGAAATTCACATTTGACTATACG 1486
Db 826 TGTGATGAGATTTTGACTATTTCTTATCTCTATAACTTTAAATGGAATCTTGATGAAT 767
Qy 1487 TACAGCGGTTCCCTCTTTTAT 1507
Db 766 TACTGAGGTATTTTTCATTAT 746
```

RESULT 4

```
US-10-511-538-86
; Sequence 86, Application US/10511538
```

```

; Publication No. US20060026700A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: TISSUE SPECIFIC GENES AND GENE CLUSTERS
; FILE REFERENCE: 16U 200 PCT
; CURRENT APPLICATION NUMBER: US/10/511,538
; CURRENT FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: US 60/372,669
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,882
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/424,336
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 60/374,823
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 60/376,558
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/381,366
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/403,648
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 4751
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (382)..(1929)
; OTHER INFORMATION:
; US-10-511-538-86

Query Match          2.4%; Score 39.8; DB 7; Length 4751;
Best Local Similarity 48.9%; Pred. No. 4.3; Indels 0; Gaps 0;
Matches 107; Conservative 0; Mismatches 112;

QY      1411 ATTGTAAATGCCTTAAGATGAGGATATCAGGAAATTTCTTCGTTTTGGGTGCAATT 1470
Db      4533 ACTTTGGAATTTCTGTAGTAGAGAAATCAGAAGACATTATCTAACTTTGTAGATACACTGT 4592

QY      1471 CCACATTGACTATACGTACAGCGGGTTCCTTTCTTATGCTATTACGTGTTATACCAATTC 1530
Db      4593 ATGATTTGGGCTTTTGTGTTCAGATTGTAATTTCTAATATAGATGAAATATTTATGCTAATA 4652

QY      1531 GTTTTACAGATTCTCTCTGAACTCGTTGGAGTGGGAATCGAGCTTTTATACAAATAGAT 1590
Db      4653 TTTTCTTATTTCAAAGCAAAATAAATGAATTTATTTGCTGTGTATAAAAAATAAAAAA 4712

QY      1591 ACTCAGTTTTTGAATGTTTTTAAAAAATAAAAAAATAAAAAA 1629
Db      4713 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4751

RESULT 5
US-09-925-065A-685225/c
; Sequence 685225, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16

```

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852195
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-852195

Query Match      2.4%; Score 38.8; DB 6; Length 622;
Best Local Similarity 63.0%; Pred. No. 3;
Matches 58; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

Qy 1535 ACAGATCTCTCTGGAAGTCTGGAAGTGGAGTGGAGTTTATACAAATACATCTC 1594
Db 52 AAAAACTCTCTTAAATTTCAATGTTAGAGGAAAGAGTAAACAAGATAATATACTC 111

Qy 1595 AGTTTGTGAATCGTTTAAAAA-AAAAAAAAAAAA 1626
Db 112 TATATTAGTATATTTTAA-AAAAAAAAAAAA 143

RESULT 8
US-11-098-686-8738/c
; Sequence 8738, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098.686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8738
; LENGTH: 194553
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8738

Query Match      2.4%; Score 38.8; DB 12; Length 194553;
Best Local Similarity 48.6%; Pred. No. 48;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 1411 ATTCTGAATGCCITTAAGATGAGGATATACAGGAATTTCTTCGTTTGGTTCGAATT 1470
Db 70218 ATTTTATATACCAATTATATTTATTTGTTATCTATTTTGTCTATTTTCTTTT 70159

Qy 1471 CCACCTTTGACATACGTACAGCGGTCCTCTTTATGCTATTACGTGTATACCAATC 1530
Db 70158 TGATTATATCTTTAGTTATTAAGTGAATCTATTTTATTTTCTTTATTTCCATTA 70099

Qy 1531 GTTTACAGATCTCTCTGGAAGTCTGTTGGAAGTGGAGTGGAGTTTATACAAATAGAT 1590
Db 70098 GAATTCATATCTCTTACATAGGGAAGATATGTTCAATATAGAAATTTTCTTAAATATA 70039
```

```
Qy 1591 ACTCAGTTTGAATGGTTTAAAAA-AAAAAAAAAAAA 1628
Db 70038 GAAAAGCTATTATGAAGTTAACTATATATCTACTATAA 70001

RESULT 9
US-09-925-065A-335502/c
; Sequence 335502, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 335502
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-335502

Query Match      2.4%; Score 38.6; DB 6; Length 569;
Best Local Similarity 63.4%; Pred. No. 3.3;
Matches 59; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1258 TGGAAATTTAGACGACATGGAGACAGACAGAAAGATTTGCAAGAGTGAACATGGG 1317
Db 224 TTGTAGACTTGGACAAGATGGTTTACAGAAAGGATATTAAGAGTGTACTGGATCTGT 165

Qy 1318 AAGCTTCCTCCACGATATTTATTTTCCCAAGCA 1350
Db 164 TATCTGTATTTTGATATGTAATGTAATCCCAAGGA 132

RESULT 10
US-09-925-065A-687471/c
; Sequence 687471, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 687471
; LENGTH: 661
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-925-065A-687471

Query Match      2.3%; Score 38; DB 6; Length 661;
Best Local Similarity 48.6%; Pred. No. 5.1; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Qy 1391 GAGTGTGATTCCTTGGATGCGATTTGTAATGCCCTTAAGATGAGGATATCAGGAAATTTTC 1450
Db 230 GTTGTGTTTATGTTTTCGCAATGTTTAGACATTTAAATATTTATATAGTCAATATTT 171

Qy 1451 TTGTTTTTTGGTTGCAATTCACATTTGACATACGATACAGCGGTTCTTTTATGCT 1510
Db 170 TCATGTTTTTTTCATGTAATGTCCTTCTCAGGATGCTCAGAGAGACCTTCCCAATGTA 111

Qy 1511 ATTACGTGTATATACCATTCGTTTACAGATTTCTTCTGTAAGTCTGTTGGAAGTGGGAATG 1570
Db 110 AGTCGTTTCTCTCTCATTTCTTTAGAGCATCATTTGTAATCTACCTGGAATTCATGTTT 51

Qy 1571 GAGGTTTATACAAATAAGATACTCAGTTTGAAT 1604
Db 50 GTATTTTGGVAGATTATGCGCTCACAGTATTATT 17
```

```
RESULT 11
US-09-925-065A-159939
; Sequence 159939, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159939
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-159939
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 535;
Best Local Similarity 57.8%; Pred. No. 5.9;
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1513 TAGGTGTATACCATTCGTTTACAGATTTCTTGTAAGTCTGTTGGAAGTGGGAATGGA 1572
Db 368 TAGTGTGTTGTTTCTTCTGTCGTAATGCAATGTACACTTTGGAAATGTTGTGTC 427

Qy 1573 GGTTTATACAAATAAGATACTCAGTTTGAATGTTTAAAAAATAAAAAAAAAAAAA 1628
Db 428 TGGTTTAATAAATTTTATGTTGTCATATTTTAAAGTAATAATAAATAACA 483
```

```
RESULT 12
US-09-925-065A-827920/c
; Sequence 827920, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
```

```
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 827920
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-827920
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 615;
Best Local Similarity 63.0%; Pred. No. 6.3;
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1535 ACAGATTCCTCTGTAAGTCTGTTGGAAGTGGGAATGAGGTTTATACAAATAAGATACTC 1594
Db 565 AAAAATCTCTTAAATTTCAATGTTAGAGGAAAGAGTAATACAAAGATAATATACTC 506

Qy 1595 AGTTTGAATGTTTAAAAAATAAAAAAAAAAAAA 1626
Db 505 TATATTTAGTGTATTTTAAAAAATAAAAAA 474
```

```
RESULT 13
US-09-925-065A-852196
; Sequence 852196, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852196
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-852196
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 622;
Best Local Similarity 63.0%; Pred. No. 6.3;
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1535 ACAGATTCCTCTGTAAGTCTGTTGGAAGTGGGAATGAGGTTTATACAAATAAGATACTC 1594
Db 52 AAAAATCTCTTAAATTTCAATGTTAGAGGAAAGAGTAATACAAAGATAATATACTC 111

Qy 1595 AGTTTGAATGTTTAAAAAATAAAAAAAAAAAAA 1626
```

```
Db      112 TATATTAGTGTATTTTAAAAAATAAAA 143
|||||
Query Match      2.3%; Score 36.8; DB 6; Length 459;
Best Local Similarity 56.7%; Pred. No. 9;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

RESULT 14
US-09-925-065A-466709/c
; Sequence 466709, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 466709
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-466709

Query Match      2.3%; Score 37; DB 6; Length 519;
Best Local Similarity 67.5%; Pred. No. 8.5;
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy      1553 TCGTGGAGTGGGAGGTTTATACAAATAGATCTCAGTTTGAATGGTTTAA 1612
Db      450 TAGTCTGAGGATGGGTGGTCTCTTACAAATAGGAACCCGCTCATCGCCTTAAA 391

Qy      1613 AAAAAAAAAAAAAAAAAA 1629
Db      390 AAAAAAAAAAAAAAAAAA 374

RESULT 15
US-09-925-065A-192552
; Sequence 192552, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192552
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-925-065A-192552

Query Match 2.3%; Score 36.8; DB 6; Length 459;
Best Local Similarity 56.7%; Pred. No. 9;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1510 TATTACGTGTATATACCATTCGTTTACAGATTCTTCTGTGAACCTCGTTGGAAGTGGGAAT 1569
Db 339 TATTCTCTGTGTTTCCCCCTTTTCCACTGTGTCTTCTATTTTCCATGCAGCAAGTAAAA 398
Qy 1570 GGAGGTTTATACAAATAAGATACACTCAGTTTGAATGGTTTAAAAA 1629
Db 399 AATTTTAAATGAAAAATCAGATTGTGAATGTTTCTCTGTAATAA 458

Search completed: February 21, 2006, 02:35:12
Job time : 541 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 20, 2006, 23:16:16 ; Search time 427 Seconds
(without alignments)

2316.700 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPSTTPSSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh

-O=/abs/ABSSWEB_spool/US10731525/runat.17022006.164911.411/app_query.fasta.1

-DB=Published Applications_NA_New -QWMT=fascap -SUFFIX=p2n.rnpbn -MINMATCH=0.1

-LOOPCU=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abes03h

-USER=US10731525 -CGN 1.1.220 @runat.17022006.164911.411 -NCPU=6 -ICPU=3

-NO_WMAP -NEG_SCORES=0 -WAIT -DISPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:

2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:

3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:

4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:

5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:

6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:

7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:

8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:

9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:

10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:

11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:

12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1745.5	71.4	1082144	12	US-11-117-187-211 Sequence 211, App
2	307.5	12.6	4055	8	US-10-793-626-3357 Sequence 3357, App
3	250.5	10.2	468	8	US-10-793-626-307 Sequence 307, App
4	170	7.0	299	12	US-11-128-061-705 Sequence 705, App

5	170	7.0	299	12	US-11-128-061-4347	Sequence 4347, App
6	170	7.0	299	12	US-11-128-049-705	Sequence 705, App
7	170	7.0	299	12	US-11-128-049-4347	Sequence 4347, App
8	159.5	6.5	1938	8	US-10-821-234-557	Sequence 557, App
9	139.5	5.7	4040	8	US-10-131-826A-425	Sequence 425, App
10	131	5.4	3155	8	US-10-793-626-3424	Sequence 3424, App
11	129.5	5.3	2578	12	US-11-136-527-219	Sequence 219, App
12	127.5	5.2	1518	8	US-10-821-234-435	Sequence 435, App
13	126	5.2	2514	12	US-11-122-329-23	Sequence 23, Appl
14	123.5	5.1	5801	12	US-11-000-463-580	Sequence 580, App
15	118	4.8	1323	8	US-10-821-234-430	Sequence 430, App
16	118	4.8	2344	8	US-10-947-249-126	Sequence 126, App
17	116	4.7	2206	12	US-11-128-061-554	Sequence 554, App
18	116	4.7	2206	12	US-11-128-049-554	Sequence 554, App
19	110.5	4.5	3133	12	US-11-000-463-108	Sequence 108, App
20	109	4.5	1881	8	US-10-467-657-5431	Sequence 5431, App
21	107	4.4	7231	12	US-11-136-527-2622	Sequence 2622, App
22	106.5	4.4	1400	12	US-11-136-527-6514	Sequence 6514, App
23	106.5	4.4	1896	12	US-11-136-527-2418	Sequence 2418, App
24	104	4.3	2712	8	US-10-821-234-8	Sequence 8, Appli
25	99	4.0	4119	12	US-11-128-061-1127	Sequence 1127, App
26	99	4.0	4119	12	US-11-128-049-1127	Sequence 1127, App
27	99	4.0	4170	12	US-11-024-959-92	Sequence 92, Appl
28	94	3.8	991	8	US-10-750-185-30544	Sequence 30544, A
29	94	3.8	991	8	US-10-750-623-30544	Sequence 30544, A
30	92.5	3.8	7659	12	US-11-091-883-498	Sequence 498, App
31	92.5	3.8	11151	12	US-11-052-554A-525	Sequence 525, App
32	92	3.8	1344	8	US-10-821-234-732	Sequence 732, App
33	92	3.8	2131	8	US-10-775-169-246	Sequence 246, App
34	91.5	3.7	552	8	US-10-467-962B-5	Sequence 5, Appli
35	91.5	3.7	4874	8	US-10-995-561-94	Sequence 94, Appl
36	90.5	3.7	2346	6	US-09-925-065A-1321	Sequence 1321, App
37	90.5	3.7	4089	8	US-10-775-169-137	Sequence 137, App
38	90.5	3.7	5053	8	US-10-995-561-93	Sequence 93, Appl
39	90.5	3.7	137671	12	US-11-121-086-47	Sequence 47, Appl
40	90	3.7	495	6	US-09-925-065A-66769	Sequence 66769, A
41	89.5	3.7	2277	12	US-11-059-275-1	Sequence 1, Appli
42	89.5	3.7	2283	8	US-10-858-730-147	Sequence 147, App
43	89.5	3.7	5676	12	US-11-134-795-5	Sequence 5, Appli
44	89.5	3.7	6221	12	US-11-134-795-6	Sequence 6, Appli
45	89	3.6	5534	12	US-11-040-472-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-11-117-187-211
; Sequence 211, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVEN, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ IDS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-11-117-187-211

Alignment Scores:
Pred. No.: 4.12e-154 Length: 1082144
Score: 1745.50 Matches: 342
Percent Similarity: 80.5% Conservatives: 43
Best Local Similarity: 71.5% Mismatches: 47

Qy	156	GlulysHisTyrGlyIleArgIleGluTyrMetPheProAsp---	AlaValGluValGln	174
Db	2801	GATAAATATCGCAATATACGCATTAAATGAAATAACCGGAACCTTACGTTAGAGGACAA	2742	
Qy	175	AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCys	194	
Db	2741	GGTGAGAAATATAATCTCGCTTTA-----TGGAAGAATGATCCTAAACCAATGTTGC	2691	
Qy	195	ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrrPilleThr	214	
Db	2690	TACATACGCAGAATTAACCACTAGAGACGATATATCTGCTGCTAGCTTGGATATCA	2631	
Qy	215	GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp	234	
Db	2630	GGTCTTAGACGAGCACAAATCACCA--ACACGAGCACATACAAATTTTCATTACAAGAT	2574	
Qy	235	ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrrAsnProValAla	254	
Db	2573	GAAGAATTTAAGTCACATAATAA-----GTGTGTCCTTAATCTATTGGACAGAA-----	2526	
Qy	255	AsnValLysGlyHisAspIleTrrAsnPheLeuArgThrMetAsnValProValAsnSer	274	
Db	2525	-----GNAAGAAGTATGGTCTTATACGTCATGAAGGATTTACCATATATATGAA	2478	
Qy	275	LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro	294	
Db	2477	TTACATGATCAAAATTAATCAAGTATTGGTTGTATTCCATGTACATCACCGCTATTTCAT	2418	
Qy	295	GlyGlnHisGluArgGluGlyArgGlyTrrTrrGluAspAlaLysAlaLysGluCysGly	314	
Db	2417	TCTAATGATTCACGTGCTGGTCTGT-----TGGTCCAATTTCTAGTAGAGCTGAATGTGGA	2364	
Qy	315	LeuHis	316	
Db	2363	TTACAT	2358	

RESULT 3

```

US-10-793-626-307
; Sequence 307, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 307
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-307

```

Alignment Scores:		
Pred. No.:	1.22e-14	468
Score:	250.50	54
Percent Similarity:	53.1%	Conservative: 32
Best Local Similarity:	33.3%	Mismatches: 61
Query Match:	10.2%	Indels: 15
DB:	8	Gaps: 6

US-10-731-525-8 (1-466) x US-10-793-626-307 (1-468)

[illegible]

Qy	175	AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCys	194
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	67	GGTGAGAAATATAATCCTGCTTTA	
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	195	ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr	214
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	118	TACATAGCAGAGATTAAACCACTAGACGCTATTATCTGCTGTGTAGCTTGGATATCA	177
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	215	GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp	234
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	178	GGTCTTAGACGAGCACCAATCACCA--ACACGAGCACATACAAATTCATTAAACAAGAT	234
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	235	ProAlaPheGluGlyMetAspGlyLyleGlySerLeuValLysTrpAsnProValAla	254
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	235	GAAGAATTTAAGTCAATTTAA-----GTGTGCTCCTTAATCTATTGTGACAGAA	282
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	255	AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer	274
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	283	-----GAAGAAGTATGCTCTTATATACGTGATAGGATTTACCATATATATGAA	330
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	275	LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro	294
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	331	TTACATGATCAAAATTATCCAAAGATTGGTTGTATTCCATGTACATCACCCGGTATTTGAT	390
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	295	GlyGlnHisGluArgGlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGly	314
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	391	TCTAATGATTCACGTGCTGCTGCT-----TGGTCCAAATTTCTAGTAAAGACTGAATGTGGA	444
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Qy	315	LeuHis	316
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117
Db	445	TTACAT	450
		:::	
		-----TGGAGAATGATCCTTAACCAATGTTGC	117

RESULT 4

```

US-11-128-061-705
; Sequence 705, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128.061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 705
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-061-705

```

Alignment Scores:	
Pred. No.:	3,82e-07
Score:	170.00
Percent Similarity:	56.0%
Best Local Similarity:	35.8%
Query Match:	7.0%
DB:	12
	DB:
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
	399
	39
	22
	37
	11
	3

US-10-731-525-8 (1-466) x US-11-128-061-705 (1-299)

Qy 198 LysValArgProLeuArgAlaLeuLysGlyLeuArgAlatrpIlethrGlyGlnArg 217
 : : : : : : : : : : : :
Dd 4 AACTAGAACCTTTAAGACGTGTTTAAACGGTAATCAGCTATGGTGTCGGGTATCAGA 63

```

QY 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237
Db 64 TCTGAGCAATCTGTA---AACAGGCACGATGATGACCAACTGGAGTGGATGAGCAA---117
QY 238 GluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLys 257
Db 118 -----AATCAGTTGATCAAGTTTCATCCGATCTTTTCCTGGTCG 156
QY 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAlaAsnSerLeuHisAla 277
Db 157 CTTGATGAGGTGAAGCAATACATTAAAGAATAATATTGTTTACACACATTACATGAT 216
QY 278 LysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHis 297
Db 217 AAGGGTTTCCCTAGTATAGGATGTCTCCTTGTACAAGACGAGTAAACACAGGCGGAG---273
QY 298 GluArgGlu-GlyArgTyrTrpTyr 305
Db 274 GATTAGAGCTGGCGTTGGTCTGG 298

RESULT 5
US-11-128-061-4347
; Sequence 4347, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4347
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-061-4347

Alignment Scores:
Pred. No.: 3,82e-07 Length: 299
Score: 170.00 Matches: 39
Percent Similarity: 56.0% Conservative: 22
Best Local Similarity: 35.8% Mismatches: 37
Query Match: 7.0% Indels: 11
DB: 12 Gaps: 3

US-10-731-525-8 (1-466) x US-11-128-061-4347 (1-299)

QY 198 LysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArg 217
Db 4 AACTAGACCTTTAGACGTGCTTTAAACGGTATACGACTATGGTTACCGTATCAGA 63
QY 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237
Db 64 TCTGAGCAATCTGTA---AACAGGCACGATGATGACCAACTGGAGTGGATGAGCAA---117
QY 238 GluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLys 257
Db 118 -----AATCAGTTGATCAAGTTTCATCCGATCTTTTCCTGGTCG 156
QY 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAlaAsnSerLeuHisAla 277

```

```
US-11-128-049-4347
; Sequence 4347, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounse, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCES: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4347
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Cricetus griseus
; US-11-128-049-4347

Alignment Scores:
Pred. No.: 3 82e-07 Length: 299
Score: 170.00 Matches: 39
Percent Similarity: 56.0% Conservative: 22
Best Local Similarity: 35.8% Mismatches: 37
Query Match: 7.0% Indels: 11
DB: 12 Gaps: 3

US-10-731-525-8 (1-466) x US-11-128-049-4347 (1-299)
QY 198 LysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArg 217
Db 4 AACTAGAACCTTTAAGACGCTGCTTTAAACGGTAATCAGCTATGGGTACCGGTATCAGA 63

QY 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237
Db 64 TCTGAGCAATCTGTA---AACAGCAGCATATACCAACCTGAGTGGATGAGCAA--- 117

QY 238 GluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaLeuValLys 257
Db 118 -----AATCAGTTGATCAAGTTTCATCCGATCTTTTCTGGTCG 156

QY 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAla 277
Db 157 CTGTATGAGGTGAAGGAATACATTAAGAAGAAATATTTTACACACATTTACATGAT 216

QY 278 LysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHis 297
Db 217 AAGGGTTTCCTAGTATAGATGCTCTCTGTGTAAGAGCAGTAAACACAGGGCGAG--- 273

QY 298 GluArgGlu-GlyArgTrpTrp 305
Db 274 GATTTAGAGCTGCGGTGGTTGG 298

RESULT 8
US-10-821-234-557
; Sequence 557, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
```

```
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC_SEQ_genes Version 1.0
; SEQ ID NO 557
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-821-234-557

Alignment Scores:
Pred. No.: 3 17e-05 Length: 1938
Score: 159.50 Matches: 112
Percent Similarity: 33.7% Conservative: 81
Best Local Similarity: 19.6% Mismatches: 206
Query Match: 6.5% Indels: 174
DB: 8 Gaps: 23

US-10-731-525-8 (1-466) x US-10-821-234-557 (1-1938)
QY 7 SerSerIleSerAlaProThrSerThrPheProSerSerGluProLysLeuProGlnIle 26
Db 352 GCCAAGATCGATGCAACCTCAGCGTCTGTGCTGCCAGCAGGTTTGATGTGATGGCTAC 411

QY 27 GlySerIleArgIleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGln 46
Db 412 CCCACCATCAAGATCCTTAAGAAG-----GGCAGCGCTGTAGACTACGAGGGCTCAGA 465

QY 47 ArgArgSerLeuValLysProValAsnAlaGluProProArgLysAspSerIleValPro 66
Db 466 ACCAGGAAGAAATTTGTCACAGGTGTCAGAGAGTCTCCAGCGCCGACTGAGCGCTCCA 525

QY 67 LeuAlaAlaThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGln 86
Db 526 CCAGAAAGTCACGCTGTGTGTGACC-----AAAGAGAACTTTGATGAA 567

QY 87 IleAlaSerAspLeuAsp----- 92
Db 568 GTTGTGATGATGCAGATATCATCTTGTGTGGAGTTTTTATGCCCATGGTGTGCACACTGC 627

QY 93 -----AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 628 AAGAAACTTGCCTCCCGAGTATGAGAAGGCGCCAGGAGCTCAGAACGCTTCTCTCTCCA 687

QY 111 IleAlaIleAla-----PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeu 129
Db 688 ATTCCCTCGCAAGGTGCGCCAGCCAGCAAAACAGACTGCGCCAGAGGTTTGTATGTC 747

QY 130 ThrGlyArgPro---PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThr 148
Db 748 TCTGGTATCCCACTCCGAAATTTTCCGC-----AAAGGAAGG-----CCTTATGAC 795

QY 149 TyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhe--- 167
Db 796 TAC-----AACGGCCACGAGAAAAATATGGAATC---GTTGATTACATGATCGAG 843

QY 168 -----ProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184
Db 844 CAGTCCGGGCTCTCCCTCAAGGAGATTCGACCCCTG-----AAGCAGGTCCAGGAG 894

QY 185 PheTyrGluAspGlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArg 204
Db 895 TTCCTGAAGATGGAGACCAT----- 915

QY 205 AlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThr 224
Db 916 -----GTCATCATCATCGGGGTCTTTAAGGGGGGAGACT----- 948

QY 225 ArgSerGluIleProValGlnValAspProAlaPheGluGlyMetAspGlyGlyIle 244
Db 949 -----GACCCAGCTTACCAGCAATACCAGGATGCCGCT 981
```

```
Qy 245 GlySerLeuVallystrpAsnProValAlaAsnVallyserGlyHisaspIletrpAsnPhe 264
Db 982 AACAACTGAGAGAAGATTACAATTTCCACACACTTTCAGACACAGAAATAGCAAAATTC 1041
Qy 265 LeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer----- 282
Db 1042 TTGAAGTCTCCAGGGCAGTGGTGTATGATGAGCCTGAGAAATTCAGTCCAGTAT 1101
Qy 283 -----Ile-GlyCysGluProCysThrArgPro----- 291
Db 1102 GAGCCCGGAGCCACATGATGAGCTCCAGGGTCCACCCAGGACTCGGCCATCAAGGAC 1161
Qy 291 ----- 291
Db 1162 TTCTGTGTAAGTACGCCCTGCGCTGGTGGCCACCGCAAGGTGTCAACGATGCTAAG 1221
Qy 292 -ValLeuProGlnHisGluArgGluGlyArgTyrTip-----TrpG1 306
Db 1222 CGCTACACAGGC-----GCCCTGTGTGTCTACTACTACGTGTGGAC 1266
Qy 306 uAspAlaLysGluCysGlyLeuHisLysGlyAsnVallyserGln----- 322
Db 1267 TTCAGCTTTGATTACAGAGCTGCAACTCAGTTTGGCGGAGCAAGTCTTAGAGGTGCC 1326
Qy 323 -----GlnLysGluAspValAsnGlyAsn-- 331
Db 1327 AAGACTTCCCTGAGTACACTTTGCCATTGCGGAC-GAAGAGGACTATGCTGGGAGGT 1385
Qy 332 -----GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrVal----- 345
Db 1386 GAAGGACTGGGCTCAGCAGAGTGGGGAGATGTAATGCCGCCATCTCGGACGAGAG 1445
Qy 346 -----ProAspIlePheAsnSerProAsn-----Va 354
Db 1446 TGGGAAGAAGTTCGCCATGAGCAGAGTGGGAGATTTGACTCTGACACCCCTCGCGAGTTGT 1505
Qy 354 lValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu----- 369
Db 1506 CACTGCTTTCAAAAAGAAAGAACTGAAGCCAGTCAATATCCAGCCAGTGCCTCAAGAA 1565
Qy 370 -----AspAr 371
Db 1566 CAACAAGGGACCGCTCAAGTCTGGTGGGAAGACCTTTGACTCCATTGTGTGACCC 1625
Qy 371 gLysGluProTrpLeuValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetG1 391
Db 1626 CAAGAAGGAGCTCTCATCGAGTCTACGACCATGTTGGGGGCACTGCAAGCAGCTAGA 1685
Qy 391 uGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLy 411
Db 1686 GCCGTGTACAACAGCTGGCCCAAGAAGTACAAGGGCCAAAGGGCCCTGTGTCATCGCAA 1745
Qy 411 sPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPh 431
Db 1746 GATG-----GAGCCACTGCCACAGCAGTCCCGCAGCGACCGCTATAAGTGGAGGCTT 1799
Qy 431 eProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerG1 451
Db 1800 CCCCACCATCTACTTCGCCCCCAGTGGGGGCAAAAAGAACCCAGTTAAATTTGAGGGTGG 1859
Qy 451 uLysArgAspValAspSerLeuMetAlaPheVal 462
Db 1860 AGACAGAGATCTGGAGCATTTTGACAAAGTTTATA 1893
```

RESULT 9

```
US-10-131-826A-425
; Sequence 425, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Deanoyers, Luc
```

```
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F3330R1C128
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US/10/131,826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 425
; LENGTH: 4040
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-131-826A-425

Alignment Scores:
Pred. No.: 0.00599 Length: 4040
Score: 139.50 Matches: 113
Percent Similarity: 37.6% Conservative: 68
Best Local Similarity: 23.5% Mismatches: 209
Query Match: 5.7% Indels: 92
DB: 8 Gaps: 24

US-10-731-525-8 (1-466) x US-10-131-826A-425 (1-4040)
Qy 27 GlySerIleArgIleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGln 46
Db 932 GGGTTACTTCGAAT-----GGAGCTGTAACTGTGTGTATGATGATAGA 973
Qy 47 ArgArgSerLeuVallyserProValAsnAlaGluProProArgLysAspSerIleValPro 66
Db 974 ATGCTTTGCCGAATGAAGAGGATCAACAGCTATCCC----- 1009
Qy 67 LeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLysGluIuAspPheGluGln 86
Db 1010 -----AGTCTCTTCATTTTTCGTCCTGG-AATGGCCCCAGTGAATATCATGGAGACAG 1062
Qy 87 IleAla-----SerAspLeuAspAsnAlaSerProLeuGluIleMetAspArgAla 103
Db 1063 ATCAAGGAGAGGTTTGTAGTGTTCATTCGATTTTGCATTCAGCATGTTAGTAAGTACAGTGACAGAACT 1122
Qy 104 LeuAsp---Lys-PheGlyAsnAspIleAlaIleAlaPheSer----- 116
```


ORGANISM: Rattus norvegicus
US-11-136-527-219

Alignment Scores:

Pred. No.: 0.0328 Length: 2578
Score: 129.50 Matches: 114
Percent Similarity: 33.5% Conservatives: 77
Best Local Similarity: 20.0% Mismatches: 211
Query Match: 5.3% Indels: 170
DB: 12 Gaps: 23

US-10-731-525-8 (1-466) x US-11-136-527-219 (1-2578)

```

Qy 7 SerSerileSerAlaProThrSerThrPheProSerGluProLysLeuProGlnIle 26
Db 634 GCAAGATGACGCGACTCAGCATCCATGTTGGCCAGCAATTCGATGTGAGTGGCTAC 693
Qy 27 GlySerileArgileSerGluAtrgProIleGlyAlaValAanPheAan----- 43
Db 694 CCCACCATCAAGATCCTGAAGAG-----GGACAGGCTGTGTGACTATGACGGCTCCAGG 747
Qy 44 -----LeuSerGlnArgArgSerLeuValLysProValAanAlaGluPro 58
Db 748 ACCCAGGAAGAAATGTTGCCAAGTCAGAGAAGTTTCCAGCCTGATGGACACCTCCA 807
Qy 59 ProArgLys-----AspSerileValProLeuAlaAla 69
Db 808 CCTGAAGTCACTTACATTGACTAAAGAGAACTTTGACGATGTGTAAATAATGACAGAC 867
Qy 70 ThrThrileVal-----AlaSerAlaSerGluThrLysGluGluAsp 83
Db 868 ATCATTTTGGTGGAGTTCTATGCCCATCGTGGCGGACACTGCAGAAACTTGGCCCTGAG 927
Qy 84 PheGluGlnIleAlaSerAspLeuAspAanAlaSerProLeuGluIleMetAspArgAla 103
Db 928 TATCAGAAGGCTGCCAAGSSWSAGCAAGCGCTCTCCCAATTCCTTAGCCCAAGATT 987
Qy 104 LeuAspLysPheGlyAanAspIleAlaIleAlaPheSerGlyAlaGluAspValAlaLeu 123
Db 988 GATCCACCGNACAGACAGACCTGGCTAAGAGTTT-----GATGCTCTGGC 1035
Qy 124 IleGluTyrAlaLysLeu-----ThrGlyArgProPheArgValPheSerLeuAspThr 141
Db 1036 TACCCCACTTTGAAATATTCGCAAGAGGAGGCTTTTGAC----- 1077
Qy 142 GlyArgLeuAanProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyIle 161
Db 1078 -----TACAATGGTCCACGAGAGAAATATGGAATT 1107
Qy 162 ArgIleGluTyrMetPhe-----ProAspAlaValGluValGlnAlaLeu-Va 177
Db 1108 ---GTTGACTACATGTTGAGCAGTCTGGACCTCTCTTCAAGGAGATTCTGACCCCTGAAG 1164
Qy 177 laArgSerLys-----GlyLeuPheSerPheTyrGl 187
Db 1165 CAGTCCAGGAGTTTCTGAAGGATGGAGATGATGTAGTCATCTCTGGAGTCTTTGAGGA 1224
Qy 187 uAspGlyHisGlnGluCysArgValArgLys---ValArgProLeuArg----- 204
Db 1225 GTCGGWGACCCAGCGCTACTTGCAGTATCAGGATGTCTGTAAACACCTTGAGAGAAGATTAC 1284
Qy 205 -----AlaLeuLysGlyLeuArgAla----- 211
Db 1285 AAATTTACCACACTTTCAGCAGCTGAATAGCTAGTTCTTCTGAAAGTCTCCTCGGGGAA 1344
Qy 212 -TriPheThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVa 231
Db 1345 CTGGTGCTGATGCAGCCTGARAAGTTCCAGTCCAAGTAT-GAGCCAGGATGCATGTGAT 1403
Qy 231 lGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAs 251
Db 1404 GGATGTTTCAGGGCTCCACAGAG----- 1425

```

```

Qy 251 nProValAlaAanValLysGlyHisAspIleTrpAanPheLeuArgThrMetAanValPr 271
Db 1426 ----GCATCAGCCATTAAAGCATGTGTGGTGAACATGACTTGGCCCTAGTGGCCACAG 1481
Qy 271 oValAanSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPr 291
Db 1482 GAAGACTTCTTAACGATGCTAAGCGGTAC-----AGCAAGCGCCCC 1520
Qy 291 oValLeuProGlyGlnHisGlu-----ArgGluGlyArgTyrTr 304
Db 1521 CCTGGTGGTGTATATACAGCGTGACTTCAGCTTTGATTACAGAACTGCTACTCAGTT 1580
Qy 304 pTTrp----- 305
Db 1581 TTGGCGTAAACAAGTCTCAGAGGTGGCCAAGGACTTCCCTGAGTACACATTGCGCATTCG 1640
Qy 306 -----GluAsp-----AlaLysAlaLysGluCysGlyLeuHisLysGlyAanValLy 321
Db 1641 TGATGAAGAAGACTATGCCACAGAGGTGAAGGACCTGGGACTG----- 1683
Qy 321 sGlnGlnLysGluGluAspValAanGlyAanGlyLeuSerGlnSer----- 336
Db 1684 -AGTGAGAGTGGAGAGACGTCAATGACGCGCATCTTAGATGAGAGTGGGAAGAAGTTTGC 1742
Qy 337 -----HisAlaAanGlyAspAla----- 342
Db 1743 CATGGAGCCAGAGGAGTTTGCATTTCAGATGCTCTCCGAGAATTGTGTCGCTTTCAAAA 1802
Qy 343 -----ThrThrValProAspIlePheAanSerProAs 353
Db 1803 AGGAAACTAAAGCCAGTTATCAATCACAGCCAGTTTCCCAAGAACCAACAGGACCAGT 1862
Qy 353 nValValAanLeuSerArgThrGlyIleGluAanLeuAlaLysLeuGluAspAsGlysgl 373
Db 1863 CAGGCTGGTGTGGGGAAGACCTTTGAT-----GCCATTGTGATGGACCCCAAAA 1913
Qy 373 uProTrpLeuValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGluSe 393
Db 1914 GGACGCTCTCATGAAATTCATGCACCGTGTGTGGGCACTGMRRCGACGTGGAGCCTGT 1973
Qy 393 rTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheAr 413
Db 1974 CTACACACGCTAGGCAAGAAGTACAAAGGGCCAGAGGACTTGTGTCATCGCCAAGATG-- 2031
Qy 413 gAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProTh 433
Db 2032 ----GATGCTACTGCCAATGACATCACCAAGCCGATACAGSWGAGGGCTTCCCCAC 2087
Qy 433 rIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLysAr 453
Db 2088 CATCTATTTTGGCCCGCAGTGGGACAAAAGAACCCCAATTAAGTTTGAGGGGTGGCAACAG 2147
Qy 453 gAspValAspSerLeuMetAlaPheValAan 463
Db 2148 AGATYTCGAGCATTTTGAGCAAGTTTATCGAT 2178

```

RESULT 12

US-10-821-234-435
; Sequence 435, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0

```
; SEQ ID NO 435
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-435

Alignment Scores:
Pred. No.: 0.0281 Length: 1518
Score: 127.50 Matches: 95
Percent Similarity: 35.6% Conservative: 65
Best Local Similarity: 21.2% Mismatches: 164
Query Match: 5.2% Indels: 125
DB: 8 Gaps: 22

US-10-731-525-8 (1-466) x US-10-821-234-435 (1-1518)

Qy 58 ProProArgLysAspSerIleValPro--LeuAlaAlaThrThrIleValAlaSerAla 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 352 CCTAGGACTGCTGATGGAATTGTCAGCCACTTGAAGAGCAGGAGCAGCAGCTTCAGTG 411
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 77 SerGluThrLysGluGluAspPheGluGluIleAlaSerAspLeuAspAsnAlaSerPro 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 412 CCTCTCAGGACTGAGGAGAAATTAAAGAAATTCATTAGTGAAGAT--GCTCTATA 468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 97 LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 469 GTAGGTTTTTCGATGATTCATTTCAGTGAGGCTCACTCCGAGTTCCCTAAAAGCAGCCAGC 528
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 117 GlyAlaGluAsp----- 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 529 AACTGTGAGGATAACTACCGATTGTCACATACGAAATGTTGAGTCTCTGTGTAACGAGTAT 588
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 -----ValAlaLeuIleGluTyrAlaLysLeuThrGlyArg----- 132
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 589 GATGATAATGAGAGGGTATCATCTTATTCGTCCTTCACATCTCACTAACAGTTTGAG 648
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 133 -----ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThr 148
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 649 GACAAGACTGTGGCATATACAGCAAAATGACCACTGGCAAAATTAATAAGTTTATC 708
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 149 TyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhePro 168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 709 CAG-----GAAACATTTTGGTATC--TGCCCTCCATCAGACAGAA 747
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 169 AspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAsp 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 748 GACATAAAGAT-----TTGATCAGGGCAGGACTTACTTATGCTTACTATGAT 798
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 189 GlyHisGlnGluCysArgValArgLysValArgProLeuArgArgAlaLeuLysGly 208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 799 GTGGACTATGAA-----AAGAACGGCTAAAGGT 825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 209 LeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIle 228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 826 TCCAACACTACTGG-----AGAAACAGGGTA 849
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 229 ProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuVal 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 850 ATGATGGTG-----GCAAGAAATTCCTGATGCTGGG-----CAC 885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 249 LysTrpAsn---ProValAlaAsnValLys-----GlyHisAspIleTrpAsnPhe--- 264
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 886 AAACCTCAACTTGTGTAGCTAGCGCAAAACCTTTAGCCATGAACCTTCGTGATTTTGGC 945
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 265 -----LeuArgThrMetAsnValProValAsnSerLeuHis---AlaLysGlyTyrVal 281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 946 TTGGAGAGCACTGCTGGAGAGATTCTCTGTTGTTGCTATCAGAACTGCTAAAGAGAGAAG 1005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 282 SerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGly 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1006 TTGTGATGAGGAGGAGTCTCGCGT----- 1032
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 302 ArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis-----LysGly 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:30:00 ; Search time 187 Seconds
(without alignments)
1094.923 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSIAPTSTPFSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2445	100.0	466	3 AAY44745	Aay44745 Soybean A
2	1951	79.8	475	8 ADX74624	Adx74624 Plant ful
3	1923	78.7	472	3 AAY44744	Aay44744 Soybean A
4	1811.5	74.1	465	3 AAG42684	Aag42684 Arabidops
5	1811.5	74.1	465	5 ABB92842	Abb92842 Herbicida
6	1811.5	74.1	465	8 ADN74399	Adn74399 Thale cre
7	1811.5	74.1	487	3 AAG42683	Aag42683 Arabidops
8	1810	74.0	463	3 AAG42685	Aag42685 Arabidops
9	1806.5	73.9	465	3 AAG17338	Aag17338 Arabidops
10	1806.5	73.9	487	3 AAG17337	Aag17337 Arabidops
11	1805	73.8	463	3 AAG17339	Aag17339 Arabidops
12	1801	73.7	458	3 AAG20363	Aag20363 Arabidops
13	1801	73.7	483	3 AAG20362	Aag20362 Arabidops
14	1797	73.5	458	3 AAG32469	Aag32469 Arabidops
15	1758.5	71.9	455	3 AAG04414	Aag04414 Arabidops
16	1738	71.1	454	3 AAG45534	Aag45534 Arabidops
17	1738	71.1	454	5 ABB91430	Abb91430 Herbicida
18	1738	71.1	454	8 ADN73975	Adn73975 Thale cre
19	1687.5	69.0	407	3 AAG04415	Aag04415 Arabidops
20	1670	68.3	406	3 AAG45535	Aag45535 Arabidops
21	1645.5	67.3	463	3 AAY44746	Aay44746 Wheat ade
22	1628.5	66.6	358	3 AAG20364	Aag20364 Arabidops
23	1624.5	66.4	358	3 AAG32470	Aag32470 Arabidops
24	1622.5	66.4	473	8 ADT58251	Adt58251 Plant poi

25	1597.5	65.3	459	4 AAG62908	Aag62908 Amino aci
26	1582.5	64.7	354	3 AAG04416	Aag04416 Arabidops
27	1573.5	64.4	354	3 AAG45536	Aag45536 Arabidops
28	1522	62.2	355	7 ABM74156	Abm74156 DNA clone
29	1308.5	53.5	292	3 AAG32471	Aag32471 Arabidops
30	1295.5	53.0	299	3 AAY44743	Aay44743 Impatiens
31	1262.5	51.6	293	3 AAY44742	Aay44742 Corn aden
32	1241.5	50.8	323	8 ADX93872	Adx93872 Plant ful
33	858.5	35.1	230	3 AAG54986	Aag54986 Arabidops
34	811.5	33.2	260	6 ADA33902	Ada33902 Acinetoba
35	745.5	30.5	269	7 AB081769	Ab081769 Pseudomon
36	731.5	29.9	235	8 ADS24762	Ads24762 Bacterial
37	540	22.1	240	9 ABM95970	Abm95970 M. xanthu
38	539	22.0	230	8 ADS30174	Ads30174 Bacterial
39	449.5	18.4	225	8 ADS30646	Ads30646 Bacterial
40	407.5	16.7	251	8 ADL05271	Adl05271 M. catarr
41	393.5	16.1	219	8 ADN27132	Adn27132 Bacterial
42	370.5	15.2	220	8 ADS21992	Ads21992 Bacterial
43	347	14.2	236	6 AAE32793	Aae32793 Mycobacte
44	325	13.3	196	8 ADS28653	Ads28653 Bacterial
45	322	13.2	247	6 AAE32794	Aae32794 Mycobacte

ALIGNMENTS

RESULT 1
AAY44745
ID AAY44745 standard; protein; 466 AA.
XX AAY44745;
AC AAY44745;
DT 04-MAY-2000 (first entry)
XX
DE Soybean APS reductase encoded by cDNA clone sl2.pk0064.g4.
XX Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;
KW sulphate assimilation protein; SAP; clone sl2.pk0064.g4.
XX
XX Glycine max.
XX
PN WO200004161-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015808.
XX
PR 14-JUL-1998; 98US-0092833P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Falco SC, Allen SM;
XX
DR WPI; 2000-182427/16.
XX N-PSDB; AA250066.
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, and in producing transgenic plants.
XX
XX Claim 1; Fig 1; 47pp; English.
XX The present sequence is soybean adenosine-5'-phosphosulfate (APS)
CC reductase, a sulphate assimilation protein (SAP). This is encoded by cDNA
CC clone designated sl2.pk0064.g4. APS reductase polypeptides are useful for
CC producing antibodies for screening cDNA expression libraries to isolate
CC full length APS reductase cDNA clones. Chimeric genes comprising the APS
CC reductase gene when transformed in a host cell can alter the level of SAP
CC expression. This can be used to increase sulphur containing compounds in
CC the cell, including the nutritionally critical amino acids cysteine and
CC methionine
XX
SQ Sequence 466 AA;

Query Match 100.0%; Score 2445; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 3.4e-243;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALAFTSSISAPTSTPSSSPKLPQIGSIRISERPIGAVNFNLSQRRSLVQKPNAPPR 60
Dy 1 MALAFTSSISAPTSTPSSSPKLPQIGSIRISERPIGAVNFNLSQRRSLVQKPNAPPR 60
Qy 61 KDSIVPLAATTIVASASETKEEDFEQIASDLNPNALMDRALDKFGNDIAFSGAED 120
Dy 61 KDSIVPLAATTIVASASETKEEDFEQIASDLNPNALMDRALDKFGNDIAFSGAED 120
Qy 121 VALIEYAKLTGRPRVPSLDTGRINPETYQLFPAVEKHGIRIYEMPPDAVEQVALVRSK 180
Dy 121 VALIEYAKLTGRPRVPSLDTGRINPETYQLFPAVEKHGIRIYEMPPDAVEQVALVRSK 180
Qy 181 GLFSFYEDGHQECRCRVRKVRPLRALKGLAWITGQRKQSPGTRSEIPVQVDPAPEGM 240
Dy 181 GLFSFYEDGHQECRCRVRKVRPLRALKGLAWITGQRKQSPGTRSEIPVQVDPAPEGM 240
Qy 241 DGGIGSLVKNPNVANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE 300
Dy 241 DGGIGSLVKNPNVANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE 300
Qy 301 GRWWEDAKAKCEGLHKGNVKQKEEDVNGLSQSHANGDATVPDIFNSPNVNLST 360
Dy 301 GRWWEDAKAKCEGLHKGNVKQKEEDVNGLSQSHANGDATVPDIFNSPNVNLST 360
Qy 361 GIENLAKLEDKPELWLVVAPWPCYQOAMEESVVDLADKLAGSTGMKVKFRADGEQKE 420
Dy 361 GIENLAKLEDKPELWLVVAPWPCYQOAMEESVVDLADKLAGSTGMKVKFRADGEQKE 420
Qy 421 FAKSELQSGFPTTLFPFKHSSRRTIKYPSEKRDVDSLMFVNALR 466
Dy 421 FAKSELQSGFPTTLFPFKHSSRRTIKYPSEKRDVDSLMFVNALR 466

RESULT 2
ADX74624
ID ADX74624 standard; protein; 475 AA.

XX AC ADX74624;
XX DT 21-APR-2005 (first entry)
XX DE Plant full length insert polypeptide seqid 43990.
XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX XX 28-APR-2003; 2003US-00425114.
XX PF 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-003985678.
XX XX (LIUJ/) LIU J.
XX XX (ZHOU/) ZHOU Y.
XX XX (KOVA/) KOVALIC D K.
XX XX (SCRE/) SCREEN S E.
XX XX (TABA/) TABASKA J E.

(CAOY/) CAO Y.
Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
WPI; 2004-180133/17.
New recombinant DNA construct, useful for improving plant tolerance to
cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
pests, for conferring increased resistance to plant disease, or for
improving yield.
Claim 1; SEQ ID NO 43990; 15pp; English.
The invention describes a recombinant DNA construct comprising a
polynucleotide consisting of a sequence encoding an amino acid sequence
available in electronic form from the US patent office at
ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
of the invention are also useful in physical arrays of molecules and as
plant breeding markers. The recombinant DNA construct is useful for
improving plant tolerance to cold, heat, drought, herbicides, extreme
osmotic conditions, pathogens or pests, for manipulating growth rate in
plant cells by modification of the cell cycle pathway, for conferring
increased resistance to plant disease, for producing galactomannan,
lignin or plant growth regulators, for increasing the rate of homologous
recombination in plants, for improving yield by modification of
photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
or by providing improved plant growth and development under at least one
stress condition or for modifying seed oil or protein yield and/or
content. This is the amino acid sequence of a plant full length insert
polypeptide that can be used in the recombinant DNA construct of the
invention.
SQ Sequence 475 AA;
Query Match 79.8%; Score 1951; DB 8; Length 475;
Best Local Similarity 80.3%; Pred. No. 4.5e-192;
Matches 382; Conservative 38; Mismatches 40; Indels 16; Gaps 7;

Qy 1 MALAF-----TSSISAPTSTF-----PSSEPKLPQIGSIRISERPIGAVNFNLSQRRSL 50
Dy 6 MALAVSTTSSSSAAAASSFFSRLLGSSSDAKAPQIGSFPPERSLVSSVVNVQRRSL 65
Qy 51 VKPNABPPRKDSIVPLAATTIVASASETKEEDFEQIASDLNPNALMDRALDKFGND 110
Dy 66 VRPLNAEPQRNDSIVPLAA-TIVAPEVEKEEDFEQIAKOLENAPLEIMDRALDKFGND 124
Qy 111 IAIAFSGAEDVALIEYAKLTGRPRVPSLDTGRINPETYQLFPAVEKHGIRIYEMPPDA 170
Dy 125 IAIAFSGAEDVALIEYAKLTGRPRVPSLDTGRINPETYQLFPAVEKHGIRIYEMPPDA 184
Qy 171 VEVOALVRSKGLSPFYEDGHQECRCRVRKVRPLRALKGLAWITGQRKQSPGTRSEIPV 230
Dy 185 VEVOALVRTKGLSFYEDGHQECRCRVRKVRPLRALKGLAWITGQRKQSPGTRSEIPV 244
Qy 231 VOVDPAPEGMDGGIGSLVKNPNVANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTR 290
Dy 245 VOVDPAPEGMDGGIGSLVKNPNVANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTR 304
Qy 291 PVLPGQHEREGRWWEDAKAKCEGLHKGNVKQKEEDVNGLSQSHANGDATVPDIFN 350
Dy 305 PVLPGQHEREGRWWEDAKAKCEGLHKGNVKQKEEDVNGLSQSHANGDATVPDIFN 361
Qy 351 SPNVNLSRTGIENLAKLEDKPELWLVVAPWPCYQOAMEESVVDLADKLAGSTGMKVG 410
Dy 362 SQDVVLSRSRGIEENLAKLEDKPELWLVVAPWPCYQOAMEESVVDLADKLAGS-GVKA 420
Qy 411 KFRADGEQKEFAKSELQSGFPTTLFPFKHSSRRTIKYPSEKRDVDSLMFVNALR 466
Dy 421 KFRADGEQKEFAKSELQSGFPTTLFPFKHSSRRTIKYPSEKRDVDSLMFVNALR 475

RESULT 3
AAY44744

Db 1 MAMSVNVSSSSSGIINSRFGVSLPEKVSQIGSLRLDLRVHVPVSNLTKGKSSSVKPL 60
 QY 55 NAEPPRRKDSIVPLAAATTVIVASATKE---EDFEQIASDLDNASPLEIMDRALDKFNDI 111
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
 QY 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHGIRIERYMPDPAV 171
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
 QY 172 EVQALVRSKGLFSFYEDGHQECRCRVRKVRPLRLALKGLRAWITGORKQOSPGRSEIPV 231
 DB 181 EVQGLVRSKGLFSFYEDGHQECRCRVRKVRPLRLALKGLRAWITGORKQOSPGRSEIPV 240
 QY 232 QVDPAPFEGMDGGISLVKWPVNVANVKGHDINWFLRTMNVVNSLHAKGVYSIGCEPCTRP 291
 DB 241 QVDPVFEGLDGGVSLVKNPVPANVEGDNVWFLRTMDVPVNTLHAAGYISIGCEPCTKA 300
 QY 292 VLPQHERBGRWWMEDAKACEGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNS 351
 DB 301 VLPQHERBGRWWMEDAKACEGLHKGNVK-ENSDDAKVNGESKS-----AVADIFKS 352
 QY 352 PNVVNLRTGIEENLAKLEDKBPWLVLVYAPWPCYCOAMEESYVDLADKLKAGSTGMKVGK 411
 DB 353 ENLVTLRSQGIENLMKLENKBPWIVLVYAPWPCYCOAMEASYDELADKLKAGS-GIKVAK 411
 QY 412 FRADGEQKEFAKQELQSGSFPTILFPFKHSRPTIKYPSEKRDVDSLMAFVNALR 466
 DB 412 FRADGQKEFAKQELQSGSFPTILFPFKNSRP- IKYPSEKRDVESLTSFLNLVR 465

RESULT 5 ABB92842

ID ABB92842 standard; protein; 465 AA.

AC ABB92842;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 2053.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.

XX Claim 5; SEQ ID NO 2053; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as
 CC herbicides

SQ Sequence 465 AA;

Query Match 74.1%; Score 1811.5; DB 5; Length 465;
 Best Local Similarity 73.9%; Pred. No. 1.2e-177;
 Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

QY 1 MALAFSTSSISAPTSTPSS-----EPKLPOIGSIRISERPIGAVNFNLS-QRRSLVKPV 54
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
 QY 55 NAEPPRRKDSIVPLAAATTVIVASATKE---EDFEQIASDLDNASPLEIMDRALDKFNDI 111
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
 QY 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHGIRIERYMPDPAV 171
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
 QY 172 EVQALVRSKGLFSFYEDGHQECRCRVRKVRPLRLALKGLRAWITGORKQOSPGRSEIPV 231
 DB 181 EVQGLVRSKGLFSFYEDGHQECRCRVRKVRPLRLALKGLRAWITGORKQOSPGRSEIPV 240
 QY 232 QVDPAPFEGMDGGISLVKWPVNVANVKGHDINWFLRTMNVVNSLHAKGVYSIGCEPCTRP 291
 DB 241 QVDPVFEGLDGGVSLVKNPVPANVEGDNVWFLRTMDVPVNTLHAAGYISIGCEPCTKA 300
 QY 292 VLPQHERBGRWWMEDAKACEGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNS 351
 DB 301 VLPQHERBGRWWMEDAKACEGLHKGNVK-ENSDDAKVNGESKS-----AVADIFKS 352
 QY 352 PNVVNLRTGIEENLAKLEDKBPWLVLVYAPWPCYCOAMEESYVDLADKLKAGSTGMKVGK 411
 DB 353 ENLVTLRSQGIENLMKLENKBPWIVLVYAPWPCYCOAMEASYDELADKLKAGS-GIKVAK 411
 QY 412 FRADGEQKEFAKQELQSGSFPTILFPFKHSRPTIKYPSEKRDVDSLMAFVNALR 466
 DB 412 FRADGQKEFAKQELQSGSFPTILFPFKNSRP- IKYPSEKRDVESLTSFLNLVR 465

RESULT 6

ADN74399

ID ADN74399 standard; protein; 465 AA.

AC ADN74399;

XX 15-JUL-2004 (first entry)

XX Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 2294.

XX plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
 KW animal feed product; thale cress; cell wall biosynthesis;
 KW nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX WPI; 2004-348466/32.

XX N-PSDB; ADN74399.

XX

PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.

XX Claim 1; SEQ ID NO 2294; 134pp; English.

XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreplication, biochemistry, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX transcription factors. This polypeptide sequence is thale cress protein
XX expressed by a gene repressed 1.3 fold or more in plants overexpressing
XX the E2Fa/Dpa transcription factor, given in an exemplification of the
XX invention.

SQ Sequence 465 AA;

Query Match 74.1%; Score 1811.5; DB 8; Length 465;

Best Local Similarity 73.9%; Pred. No. 1.2e-177;

Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSSTFPSS-----EPKLPQIGTSIRISERPIGGAVNFMNLS-QRRSLVKPV 54

Db 1 MAMSVNVSSSSSGIINSRFGVSLPEKPVQSIGSLRLDLDRVHVAVSLNSGKRSSSVKPL 60

Qy 55 NAEPRKDSIVPLAATTIVASASTKE---EDFQIASDLDNASPLEIMDRALDKFNDI 111

Db 61 NAEPTKDSIMPLAATVMAETAEVEVVEIEDFDELAKUENASPLEIMDKALEKYGNDI 120

Qy 112 AIAFSGAEDVALIEYAKLTGRPFVFSLDTGRINPETYQLFDAVEKHGIRIETMYFPDAV 171

Db 121 AIAFSGAEDVALIEYAKLTGRPFVFSLDTGRINPETYQLFDAVEKHGIRIETMYFPDVS 180

Qy 172 EVQALVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLRAWITCQRKQDQSPGTRSEIPVV 231

Db 181 EVQGLVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLKAWITCQRKQDQSPGTRSEIPVV 240

Qy 232 QVDPAFEGMDGGISLVKNPVPANVKGHDINFLRTMNPVNSLHAKGYVSI GCEPCTRP 291

Db 241 QVDPVFEGLDGGVSLVKVNPVANVEGNDVNFRTMDVPNTLHAAGYISIGCEPCTKA 300

Qy 292 VLPQHEREGRWWEADAKCEGLHGNVKQKQEDVNGLSQSHANGDATTVPDIFNS 351

Db 301 VLPQHEREGRWWEADAKCEGLHGNVK-ENSDDAKVNAGESK------AVADIFKS 352

Qy 352 PNVTNLSRTGIENLAKLEDKRPWLVLVLPWPCYQCAMEESYVDLADKLAGSTGMKVKGK 411

Db 353 ENLVTLSRQGIENLAKLENKRKEPWIVLVLPWPCFCQAMEASYDELADKLAGS-GIKVAK 411

Qy 412 FRADGEQKEFAKSEIOLGSPFTIILFFPKHSRRPPIKYPSEKRDVDSLMAFVNLAR 466

Db 412 FRADGDQKEFAKSEIOLGSPFTIILFFPKNSRRP-IKYPSEKRDVESLTSFLNLVR 465

RESULT 7

AAG42683

ID AAG42683 standard; protein; 487 AA.

XX

AC AAG42683;

XX

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53259.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-00301439.
XX 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.

PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	02-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	08-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	08-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
<hr/>					
PR	27-AUG-1999;	99US-0151066P.	PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.	PR	30-AUG-1999;	99US-0151080P.
PR	31-AUG-1999;	99US-0151438P.	PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.	PR	07-SEP-1999;	99US-0152363P.
PR	07-SEP-1999;	99US-0152363P.	PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.	PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.	PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.	PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.	PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.	PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.	PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.	PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.	PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.	PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.	PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159329P.	PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.	PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.	PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.	PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.	PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-OCT-1999;	99US-0160980P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.	PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.	PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.	PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.	PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.	PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.	PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.	PR	29-OCT-1999;	99US-0162142P.
PR	29-OCT-1999;	99US-0162142P.			
<hr/>					
Query Match 74.1%; Score 1811.5; DB 3; Length 487;					
Best Local Similarity 73.9%; Pred. No. 1.3e-177;					
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;					
Qy	1	MALAFSTTSISAPTSPTSPSS----	BKLPQIGSIRISERPIGCAVNFNLS-QRRSLVKPV	54	
Db	23	MAWSNVSSSSSGSIINSRFGVSLKPKVQISGLRLDRVHVAPVSLNSCKSSSVKPL	82		
Qy	55	NAEPPRKDSIVPLAATTIVASATKX---	EDPEQIASDLDNASPLEIMDRALDKFGNDI	111	
Db	83	NAEPTKDSMIPLAATWAEIAEEVEVEIEDFEELAKKLENASPLEIMDKALEKYGNDI	142		
Qy	112	ATAFSGAEDVALIEYAKLTGRPRVPSLDTGRINLPETYQLFDAVEKHGIRIYMPDVA	171		
Db	143	ATAFSGAEDVALIEYAKLTGRPRVPSLDTGRINLPETYQLFDAVEKHGIRIYMPDVS	202		
Qy	172	EQVALVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLRAWITGQRKDSGPTRSIPVV	231		
Db	203	EVQGLVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLRAWITGQRKDSGPTRSIPVV	262		
Qy	232	QVDPAFEGMDGGIGSLVKWNPVANVKGHDINWPLRTMNPVNSLHAKGYVSIQCEPCTRP	291		

Db 263 QVDVPVRGLDGVGSLVKMNPVANVEGNDVNMFLRTWDPVNTLHAAGYISIGCEPCTKA 322
Qy 292 VLFQOHERGRWWEDAKAECGLHKNVQOKEEDVNGLSQSHANGDATTVDPDFNS 351
Db 323 VLFQOHERGRWWEDAKAECGLHKNVK-ENSDDAKVNGESKS-----AVADIFKS 374
Qy 352 PNVNLSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVKG 411
Db 375 ENLVTLSRQGIENLMLKENRKEPWIVLYAPWCPFCQAMEASYDELADKLAGS-GIKVAK 433
Qy 412 FRADGEQKEFAKSELQSGFPTTLFFPKHSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 434 FRADGDQKEFAKSELQSGFPTTLVFPKNSRP- IKYPSEKRDVESLTSFLNLVR 487

RESULT 8
AAG42685
ID AAG42685 standard; protein; 463 AA.
AC AAG42685;
XX
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53261.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142977P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143624P.
PR 14-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.

PR 30-APR-1999; 99US-01332048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-01332486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 18-MAY-1999; 99US-0134941P.
PR 18-MAY-1999; 99US-0135124P.
PR 20-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137502P.
PR 04-JUN-1999; 99US-0137724P.
PR 07-JUN-1999; 99US-0138094P.
PR 08-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 10-JUN-1999; 99US-0139119P.
PR 14-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 16-JUN-1999; 99US-0139492P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	PN	EP1033405-A2.
PR	21-OCT-1999;	99US-0160767P.	XX	
PR	21-OCT-1999;	99US-0160768P.	PD	
PR	21-OCT-1999;	99US-0160770P.	XX	06-SEP-2000.
PR	21-OCT-1999;	99US-0160814P.	PF	
PR	21-OCT-1999;	99US-0160815P.	XX	25-FEB-2000; 2000EP-00301439.
PR	22-OCT-1999;	99US-0160980P.	XX	
PR	22-OCT-1999;	99US-0160981P.	PR	25-FEB-1999;
PR	22-OCT-1999;	99US-0160989P.	PR	05-MAR-1999;
PR	25-OCT-1999;	99US-0161404P.	PR	09-MAR-1999;
PR	25-OCT-1999;	99US-0161405P.	PR	23-MAR-1999;
PR	25-OCT-1999;	99US-0161406P.	PR	25-MAR-1999;
PR	25-OCT-1999;	99US-0161359P.	PR	99US-0126264P.
PR	26-OCT-1999;	99US-0161360P.	PR	99US-0126785P.
PR	26-OCT-1999;	99US-0161361P.	PR	99US-0127462P.
PR	28-OCT-1999;	99US-0161920P.	PR	99US-0128234P.
PR	28-OCT-1999;	99US-0161922P.	PR	99US-0128714P.
PR	28-OCT-1999;	99US-0161933P.	PR	99US-0129845P.
PR	29-OCT-1999;	99US-0162142P.	PR	99US-0130077P.
Query Match			PR	99US-0130449P.
Best Local Similarity 73.7%; Pred. No. 3.9e-177;			PR	99US-0130510P.
Matches 350; Conservative 48; Mismatches 58; Indels 19; Gaps 7;			PR	99US-0130891P.
QY	1	MALAPTSSISAPTSTPSS-----EPKLPQIGSIRISERPIGGAVNPNLS-QRRSLVKPV 54	PR	99US-0131449P.
Db	1	MAMSVNVSSSSSGIINSRGVSLPKVSIQISLRLLDRVHVAPVSLNSGKRSSSVKPL 60	PR	99US-0132048P.
QY	55	NASPPRKDSIVPLAATTIVASSETKE--EDPEQIASDLDNASPLEIMDRALDKGNDI 111	PR	99US-0132407P.
Db	61	NABPKTNDMSIPLAATMVABIEVEVEIEDPEELAKLENASPLEIMDKALEKYGNDI 120	PR	99US-0132484P.
QY	112	AIAPGAEDVALLEYAKLTGRPRVFSLDTRGLNPETYQLFDAVEKHGIRIETMPDDAV 171	PR	99US-0132485P.
Db	121	AIAPGAEDVALLEYAHLTGPRVFSLDTRGLNPETYRFDVAVEKHGIRIETMPDVS 180	PR	99US-0132486P.
QY	172	EVQALVRSKGLFYEDGHQECRCRVRKVRPLRALKGLRAMITGQRKQSPGTRSEIPVV 231	PR	99US-0132487P.
Db	181	EVQGLVRSKGLFYEDGHQECRCRVRKVRPLRRALKGLKAWITGQRKQSPGTRSEIPVV 240	PR	99US-0132863P.
QY	232	QVDPAGEMDGGIGSLVKNPVPANVKGHDINWFLRTMNVNPNVSLHAKGYISICECTRP 291	PR	99US-0134256P.
Db	241	QVDPVFEGLDGGVSLVKNPVPANVEGDNVNFRTMDVFNVLTHAAGYISICEPCTKA 300	PR	99US-0134218P.
QY	292	VLPQCHEREGRWWEAKACEGLHKGNVKQKEEDVNGLSOSHANGDATTVPDIENS 351	PR	99US-0134219P.
Db	301	VLPQCHEREGRWWEAKACEGLHKGNVK-ENSDDAKYNAGESKS-----AVADIFKS 352	PR	99US-0134221P.
QY	352	PNVNLISRTGIENLAKLEDKRPMLVVLVYAPWCPYCOAMESYVDLADKLAGSTGMKVKG 411	PR	99US-0134370P.
Db	353	ENLVTLISRQGIENLMLKENRKEPMIVVLVYAPWCPYCOAMEASYDELADKLGS-GIKVAK 411	PR	99US-0134768P.
QY	412	FRADGQKEPAKSELQIGSPFTILFPFKHSRRTIKYPSEKRDVDSLMFVNLAR 466	PR	99US-0134941P.
Db	412	FRADGQKEPAKSELQIGSPFTILFPFKNSRP-IKYPSEKRDVESLTSFLNLVR 465	PR	99US-0135124P.
RESULT 10			PR	99US-0135242P.
AAG17337			PR	99US-0135353P.
ID	AAG17337 standard; protein; 487 AA.		PR	99US-0135629P.
XX			PR	99US-0136021P.
AC			PR	99US-0136392P.
DT	17-OCT-2000 (first entry)		PR	99US-0136782P.
XX			PR	99US-0137222P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 18318.		PR	99US-0137502P.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	99US-0138094P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	99US-0138452P.
KW	termination sequence.		PR	99US-0139453P.
XX			PR	99US-0139457P.
OS	Arabidopsis thaliana.		PR	99US-0139458P.
XX			PR	99US-0139459P.

PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149729P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.9%; Score 1806.5; DB 3; Length 487;

Best Local Similarity 73.7%; Pred. No. 4.2e-177;

Matches 350; Conservative 73; Mismatches 58; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSTRISERPPIGGAVNFNL-ORSLVKPV 54

Db 23 MAMSVNVSSSSSGIINSRFGVSLPEPKVQIGSLRLDLDRVHVAPVSLNLGKRSSSVKPL 82

Qy 55 NAEPPRKDSIVPLAATTIVASASPTKE---EDPEQIASDLDNASPLEIMDRALDKFGNDI 111

Db 83 NAEPKTNDSMIPLAANTWVAIEAEVVEVEIDFELAKLENASPLEIMDKALEKYGNDI 142

Qy 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFADVEKHYGIRIETMFPDVA 171

Db 143 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFADVEKHYGIRIETMFPDVS 202

Qy 172 EVQALVRSGKGLFSFYEDGHQECCHVRKVRPLRRALKGLRAWITGQRKDOOSPTRSEIPVV 231

Db 203 EVQGLVRSGKGLFSFYEDGHQECCHVRKVRPLRRALKGLKAWITGQRKDOOSPTRSEIPVV 262

Qy 232 QVDPAPFSGMDGGIGSLVKWNPVANVKGHDITWNFLRTWNPVNVNSLHAKGYVSGICEPCTRP 291

Db 263 QVDPVFGLDGGVGLVKWNPVANVEGNDVWNFURTDVDPVNTLHAAGYISIGCEPCTKA 322

Qy 292 VLPQHEREGRWWEWEDAKACEGLHKGKGVKQKKEEDVNGNGLSQSHANGDATTVPDIFNS 351

Db 323 VLPQHEREGRWWEWEDAKACEGLHKGKGVK-ENSDDAKAVNGESKS-----AVADIFKS 374

Qy 352 PNVNLSRTGIENLAKLEDKPEWLVLYAPWCPCQAMEESYVDLADKLKAGSTGMKVGK 411

Db 375 ENLVTLSRQGIENLMLKENRKEPWIVLYAPWCPCQAMEASYDELADKLKAGS-GIRVAK 433

Qy	412	FRADGQEKFAKSELQIGSPPTILFPKCHSRRTIKYPSEKRDVDSILMAFVNLR	466
Db	434	FRADGQEKFAKSELQIGSPPTILFPKCHSRRTIKYPSEKRDVDSILMAFVNLR	487
<p>RESULT 11</p> <p>AAG17339</p> <p>ID AAG17339 standard; protein; 463 AA.</p> <p>XX AC AAG17339;</p> <p>DT 17-OCT-2000 (first entry)</p> <p>XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 18320.</p> <p>XX Protein identification; signal transduction pathway; metabolic pathway;</p> <p>KW hybridisation assay; genetic mapping; gene expression control; promoter;</p> <p>KW termination sequence.</p> <p>XX OS Arabidopsis thaliana.</p> <p>XX EP1033405-A2.</p> <p>PN 06-SEP-2000.</p> <p>PD 25-FEB-2000; 2000EP-00301439.</p> <p>PF 25-FEB-1999; 99US-0121825P.</p> <p>XX 05-MAR-1999; 99US-0123180P.</p> <p>PR 09-MAR-1999; 99US-0123548P.</p> <p>PR 23-MAR-1999; 99US-0125788P.</p> <p>PR 23-MAR-1999; 99US-0126264P.</p> <p>PR 29-MAR-1999; 99US-0126785P.</p> <p>PR 01-APR-1999; 99US-0127462P.</p> <p>PR 06-APR-1999; 99US-0128234P.</p> <p>PR 08-APR-1999; 99US-0128714P.</p> <p>PR 16-APR-1999; 99US-0129845P.</p> <p>PR 19-APR-1999; 99US-0130077P.</p> <p>PR 21-APR-1999; 99US-0130449P.</p> <p>PR 23-APR-1999; 99US-0130510P.</p> <p>PR 23-APR-1999; 99US-0130891P.</p> <p>PR 28-APR-1999; 99US-0131449P.</p> <p>PR 30-APR-1999; 99US-0132048P.</p> <p>PR 30-APR-1999; 99US-0132407P.</p> <p>PR 04-MAY-1999; 99US-0132484P.</p> <p>PR 05-MAY-1999; 99US-0132485P.</p> <p>PR 06-MAY-1999; 99US-0132486P.</p> <p>PR 06-MAY-1999; 99US-0132487P.</p> <p>PR 07-MAY-1999; 99US-0132863P.</p> <p>PR 11-MAY-1999; 99US-0134256P.</p> <p>PR 14-MAY-1999; 99US-0134218P.</p> <p>PR 14-MAY-1999; 99US-0134219P.</p> <p>PR 14-MAY-1999; 99US-0134221P.</p> <p>PR 14-MAY-1999; 99US-0134370P.</p> <p>PR 18-MAY-1999; 99US-0134768P.</p> <p>PR 19-MAY-1999; 99US-0134941P.</p> <p>PR 20-MAY-1999; 99US-0135124P.</p> <p>PR 21-MAY-1999; 99US-0135353P.</p> <p>PR 24-MAY-1999; 99US-0135629P.</p> <p>PR 25-MAY-1999; 99US-0136021P.</p> <p>PR 27-MAY-1999; 99US-0136392P.</p> <p>PR 28-MAY-1999; 99US-0136782P.</p> <p>PR 01-JUN-1999; 99US-0137222P.</p> <p>PR 03-JUN-1999; 99US-0137528P.</p> <p>PR 04-JUN-1999; 99US-0137502P.</p> <p>PR 07-JUN-1999; 99US-0137724P.</p> <p>PR 08-JUN-1999; 99US-0138094P.</p> <p>PR 10-JUN-1999; 99US-0138540P.</p> <p>PR 10-JUN-1999; 99US-0138847P.</p> <p>PR 14-JUN-1999; 99US-0139119P.</p> <p>PR 16-JUN-1999; 99US-0139452P.</p> <p>PR 16-JUN-1999; 99US-0139453P.</p> <p>PR 17-JUN-1999; 99US-0139492P.</p>			

PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147915P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.

PR	25-OCT-1999;	99US-01611405P.	PR	25-MAR-1999;	99US-0126264P.
PR	25-OCT-1999;	99US-01611406P.	PR	29-MAR-1999;	99US-0126785P.
PR	26-OCT-1999;	99US-01611359P.	PR	01-APR-1999;	99US-0127462P.
PR	26-OCT-1999;	99US-01611360P.	PR	06-APR-1999;	99US-0128234P.
PR	26-OCT-1999;	99US-01611361P.	PR	08-APR-1999;	99US-0128714P.
PR	28-OCT-1999;	99US-01611920P.	PR	16-APR-1999;	99US-0129845P.
PR	28-OCT-1999;	99US-01611921P.	PR	19-APR-1999;	99US-0130077P.
PR	28-OCT-1999;	99US-01611933P.	PR	21-APR-1999;	99US-0130449P.
PR	28-OCT-1999;	99US-01611934P.	PR	23-APR-1999;	99US-0130510P.
PR	28-OCT-1999;	99US-01621142P.	PR	23-APR-1999;	99US-0130891P.
Query Match 73.7%; Score 1801; DB 3; Length 458;			PR	28-APR-1999;	99US-0131449P.
Best Local Similarity 74.0%; Pred. No. 1.4e-176;			PR	30-APR-1999;	99US-0132048P.
Matches 348; Conservative 51; Mismatches 55; Indels 16; Gaps 6;			PR	30-APR-1999;	99US-0132407P.
QY	1	MALAF-----TSSISAPTSPTSPSEPKLPQIGSRISRISRPPIGGAVFNLSQRRSLVKPVNA 56	PR	04-MAY-1999;	99US-01332484P.
DB	1	MALAINVSSSSSAISSSPSDDLKVTKIGSLRLNRTNVSASLSLGRKSSVKALNV 60	PR	05-MAY-1999;	99US-01332485P.
QY	57	EPPRKDSIVPLAATTIVASASSETKEEDFEQIATSDLDNASPLEIMDRALDKFGNDIALAFS 116	PR	06-MAY-1999;	99US-01332486P.
DB	61	QSITKESIV---ASEVTEKLDVVEVEDFEHLAKLENASPLEIMDKALEFGNDIALAFS 117	PR	06-MAY-1999;	99US-01332487P.
QY	117	GAEDVALIEYAKLTGRFPRVFSLDTGRINPETYQLFDVAEKHYGIRIEYMPDPDAVEYOAL 176	PR	11-MAY-1999;	99US-01332487P.
DB	118	GAEDVALIEYAKLTGRFPRVFSLDTGRINPETYQLFDVAEKHYGIRIEYMPDPDAVEYOAL 177	PR	14-MAY-1999;	99US-01332487P.
QY	177	VRSKGLFSFYEDGHCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRSEIPVQVDPDA 236	PR	14-MAY-1999;	99US-01332487P.
DB	178	VRNKGKLSFYEDGHCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRSEIPVQVDPV 237	PR	14-MAY-1999;	99US-01332487P.
QY	237	FEQDGGIGSIVKKNPVANVKKVGHDIWNFLRTMNPVNSLHAKGVVSGCEPCTRPVLPQG 296	PR	14-MAY-1999;	99US-01332487P.
DB	238	FEGLDGGVGLVKNPVANVGGNDVNFRTMDVPVNTLHAGVVSIGCEPCTRAVLPGQ 297	PR	14-MAY-1999;	99US-01332487P.
QY	297	HEREGRWWEDAKAKECGLHKGNNVKQKEEDVNGSLSQSHANGDATTPDIFNSPNVNV 356	PR	14-MAY-1999;	99US-01332487P.
DB	298	HEREGRWWEDAKAKECGLHKGNIKE-----NTNGNATANVNGTA-TVADIFNSENVNV 350	PR	14-MAY-1999;	99US-01332487P.
QY	357	LSRTGIENLAKLEDKRPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVGKFRADG 416	PR	14-MAY-1999;	99US-01332487P.
DB	351	LSRQGIENLAKLENKREAWIVLVYAPWCPYCOAMEASFDELADKLGS-GYVAKFRADG 409	PR	14-MAY-1999;	99US-01332487P.
QY	417	EOKEFAKSELQGSFPFLLRPKHSSRPTIKYPSEKRDVDSLMAFVNALR 466	PR	14-MAY-1999;	99US-01332487P.
DB	410	DQKDFAKKELQGSFPFLLVFPKNSRP-IKYPSEKRDVDSLTSFLNLVR 458	PR	14-MAY-1999;	99US-01332487P.
RESULT 13			PR	14-MAY-1999;	99US-01332487P.
ID	AAG20362	standard; protein; 483 AA.	PR	14-MAY-1999;	99US-01332487P.
AC	AAG20362;		PR	14-MAY-1999;	99US-01332487P.
XX	XX		PR	14-MAY-1999;	99US-01332487P.
DT	17-OCT-2000	(first entry)	PR	14-MAY-1999;	99US-01332487P.
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 22519.	PR	14-MAY-1999;	99US-01332487P.
XX	XX		PR	14-MAY-1999;	99US-01332487P.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	14-MAY-1999;	99US-01332487P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	14-MAY-1999;	99US-01332487P.
XX	termination sequence.		PR	14-MAY-1999;	99US-01332487P.
OS	Arabidopsis thaliana.		PR	14-MAY-1999;	99US-01332487P.
XX	XX		PR	14-MAY-1999;	99US-01332487P.
FN	EP1033405-A2.		PR	14-MAY-1999;	99US-01332487P.
PD	06-SEP-2000.		PR	14-MAY-1999;	99US-01332487P.
XX	XX		PR	14-MAY-1999;	99US-01332487P.
XX	25-FEB-2000;	2000EP-00301439.	PR	14-MAY-1999;	99US-01332487P.
PR	25-FEB-1999;	99US-0121825P.	PR	14-MAY-1999;	99US-01332487P.
PR	05-MAR-1999;	99US-0123180P.	PR	14-MAY-1999;	99US-01332487P.
PR	09-MAR-1999;	99US-0123548P.	PR	14-MAY-1999;	99US-01332487P.
PR	23-MAR-1999;	99US-0125788P.	PR	14-MAY-1999;	99US-01332487P.

```
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147152P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.7%; Score 1801; DB 3; Length 483;
Best Local Similarity 74.0%; Pred. No. 1.5e-176;
Matches 348; Conservative 51; Mismatches 55; Indels 16; Gaps 6;

Qy 1 MALAP----TSSISAPTSTPPSEPKLQIGSTRISERPIGGAVNENLRSORSLVKPVNA 56
Db 26 MALAINVSSSSSAISSSPSSDLKVTKIGSLRLNRTNVSAAISLSLGRSSVKALNV 85
Qy 57 EPPRKDSIVPLAATTIVASASETKEDFEQIASDLDONASPLEIMDRALDKFGNDIAFAS 116
Db 86 QSITKESIV---ASEVTEKLDVVEDEDFBELAKRLENASPLEIMDKALEKFGNDIAFAS 142
Qy 117 GAEDVALIEYAKLTGRPFVFSLDTGRLNPETQYOLFDAVEKHYGIRIENYMPDAVEQAL 176
Db 143 GAEDVALIEYAKLTGRPFVFSLDTGRLNPETQYOLFDAVEKHYGIRIENYMPDAVEQAL 202
Qy 177 VRSGKLFSPYEDGHQCCRRVKRPLRLRALKGLRAMITQQRKQDQSPQTRSEIPVQVDP 236
Db 203 VRNKGKLFSPYEDGHQCCRRVKRPLRLRALKGLRAMITQQRKQDQSPQTRSEIPVQVDP 262
Qy 237 FEGMDGGIGSLVKWNPVNVKGDHIDWNFLRTMVPVNSLHAKGVYSGPCPTRPVLPGQ 296
Db 263 FEGLDGGVGSVKWNPVNVKGDHIDWNFLRTMVPVNSLHAKGVYSGPCPTRPVLPGQ 322
Qy 297 HEREGRWWMEDAKAKCGLHKGKVKQKEDVNGNGLSOSHANGDATTVDPDIENSPNVN 356
Db 323 HEREGRWWMEDAKAKCGLHKGKVKQKEDVNGNGLSOSHANGDATTVDPDIENSPNVN 375
Qy 357 LSRTGIENLAKLEDKRPEMLVLVYAPWCPYQAMEESYVDLADKLAGSTGMKVGKPRADG 416
Db 376 LSRQGIENLAKLEDKRPEMLVLVYAPWCPYQAMEASFDLADKLGGSGVKVAKFRADG 434
Qy 417 EQKEPAKSELQIGSPFTILFFPKHSSRPITKYPSEKRDVDSLMAFVNALR 466
Db 435 DQKDFAKKELQIGSPFTILFFPKHSSRP-ISKYPSEKRDVDSLTSFLNLR 483

RESULT 14
AAG32469
ID AAG32469 standard; protein; 458 AA.
XX
AC AAG32469;
```


PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142309P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145227P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:33:35 ; Search time 40 Seconds
(without alignments)
1120.925 Million cell updates/sec

Title: US-10-731-525-8
Perfect score: 2445
Sequence: 1 MALAFTSSISAPSTPSSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811.5	74.1	465	2 B85058	5'-adenylylsulfate
2	1797	73.5	458	2 T49106	PRH26 protein - Ar
3	1788.5	73.1	463	2 T10065	phosphoadenylyl-su
4	1738	71.1	454	2 C96648	hypothetical prote
5	1648	67.4	406	2 S71242	phosphoadenylyl-su
6	1555.5	63.6	422	2 S71243	phosphoadenylyl-su
7	1216	49.7	423	2 T52251	5'-adenylylsulfate
8	745.5	30.5	267	2 H83426	3'-phosphoadenosin
9	519.5	21.2	239	2 D90470	hypothetical prote
10	470.5	19.2	244	2 H81905	probable phosphoad
11	463	18.9	246	2 C81111	phosphadenosine p
12	406	16.6	260	2 H97458	hypothetical prote
13	406	16.6	260	2 AB2677	phosphadenosine p
14	325	13.3	196	2 C69839	3'-phosphadenosin
15	305.5	12.5	254	2 C70682	probable cysH - My
16	302.5	12.4	231	2 F83835	3'-phosphadenosin
17	296.5	12.1	233	2 H69611	phosphadenylyl-su
18	290.5	11.9	244	2 H97388	phospho-adenylylau
19	268.5	11.0	252	2 AG3472	phosphadenylyl-su
20	259	10.6	241	2 AH2363	phosphadenosine p
21	244.5	10.0	255	2 C75594	phosphadenosine p
22	235.5	9.6	237	2 D82674	3'-phosphadenosin
23	228	9.3	239	1 S34193	phosphadenylyl-su
24	214	8.8	266	2 S67435	phosphadenylyl-su
25	213	8.7	244	2 AE0858	3'-phosphadenosin
26	211	8.6	244	1 C34354	phosphadenylyl-su
27	210	8.6	244	1 RDECPA	phosphadenylyl-su
28	210	8.6	261	2 S59826	phosphadenylyl-su
29	207	8.5	244	2 B85926	3'-phosphadenosin

30	207	8.5	244	2 A91081	3'-phosphadenosin
31	197	8.1	253	2 D82329	phosphadenosine p
32	193.5	7.9	249	2 S74657	phosphadenylyl-su
33	168.5	6.9	244	2 D84979	phosphadenosine p
34	167.5	6.9	632	2 B75215	5'-adenylylphospho
35	162.5	6.6	232	1 S28609	phosphoadenylyl-su
36	157.5	6.4	480	2 B64308	hypothetical prote
37	154.5	6.3	266	2 C96913	sulfate adenylylate
38	154	6.3	532	2 T07927	protein disulfide-
39	154	6.3	638	1 ISMSER	protein disulfide-
40	154	6.3	645	1 A23723	protein disulfide-
41	151	6.2	630	2 E71451	hypothetical prote
42	146.5	6.0	443	2 D86183	hypothetical prote
43	145	5.9	643	1 S32476	protein disulfide-
44	143	5.8	504	2 S41661	protein disulfide-
45	141	5.8	411	2 E64421	hypothetical prote

ALIGNMENTS

RESULT 1

B85058

5'-adenylylsulfate reductase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: B85058

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: B85058

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-465 <STO>

A;Cross-references: UNIPROT:P92979; UNIPARC:UPI000000BD8A; GB:NC_001268; NID:g7267219; P

C;Genetics:

A;Gene: AT4G04610

A;Map position: 4

Query Match 74.1%; Score 1811.5; DB 2; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.7e-132;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

QY	1	MALAFTSSISAPSTPSS-----EPKLPQIGSIRISERPIGAVNFNLS-QRRSLVKPV	54
DB	1	MAMSVNVSSSSSGIINRFGVLEPKVQIGSLRLDRVHVAVPSVNLNLSKSSSSVKPL	60
QY	55	NAEPPRKDSIVPLAATTIVASASETKE---EDFEQIASLDLNASPLEIMDRALDKFGNDI	111
DB	61	NAEPTKDSMIPLAATMVAIEAEVEVVEIEDFEELAKKLENASPLEIMDKALEKYGNDI	120
QY	112	ATAFSGAEVVALEIYAKLTGRPPRVPSLDTGRINLPETYQLFDAVEKHGIRIIRIYMPDDAV	171
DB	121	ATAFSGAEVVALEIYAKLTGRPPRVPSLDTGRINLPETYRFFDAVEKHGIRIIRIYMPDVS	180
QY	172	EQVALVRSKGLSFYEDGHQECRCVRKVRPLRALGLAWITGQRKQSPGFRSETPVV	231
DB	181	EQVGLVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLAWITGQRKQSPGFRSEIPVV	240
QY	232	QVDPAPFEGMDGGTIGSLVKVNPVANVKGHDIWNFLRTMNPVNSLHAKGVYVSGICEPCTRP	291
DB	241	QVDPVFEGLDGGVSLVKVNPVANVSGNDVWNFLRTMDVPVNTLHAAGYISIGICEPCTKA	300
QY	292	VLPQGHREGRWWEDAKAKGCLHKGNYKQKKEEDVNGSLGQSQSHANGDATTVDPIDFNS	351
DB	301	VLPQGHREGRWWEDAKAKGCLHKGNYK-ENSDDAKVNKESK-----AVADIFKS	352
QY	352	PNNVNLSPGNIENLAKLEDKBPWLVLVYAPCPYCOAMEESVVDLADKLAGSTGMKVGK	411
DB	353	ENLVTLVSGIENLMLKLENKBPWIVLVYAPMCPFCOAMEASYDELDADKLKAGS-GIKVAK	411
QY	412	FRADGQKFEAKSELQSGSFPTILFPFKHSRRTIKYPSEKRDVDSLMAFVNALR	466

RESULT 7
T52251
5'-adenylylsulfate reductase [EC 1.8.99.-] [validated] - green alga (Enteromorpha intest
C;Species: Enteromorpha intestinalis (hollow green seaweed)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52251
R;Gao, Y.; Schofield, O.M.; Leustek, T.
Plant Physiol. 123, 1087-1096, 2000
A;Title: Characterization of sulfate assimilation in marine algae focusing on the enzyme
A;Reference number: Z26003; MUID:20349720; PMID:10889258
A;Accession: T52251
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-423 <GAO>
A;Cross-references: UNIPROT:O81350; UNIPARC:UPI000000A273A; EMBL:AF069951; PIDN:AAC26855.
C;Function:
A;Description: catalyzes reduced glutathione-dependent reduction of adenylylsulfate to a
C;Keywords: oxidoreductase

Query Match 49.7%; Score 1216; DB 2; Length 423;
Best Local Similarity 60.1%; Pred. No. 2.1e-86;
Matches 239; Conservative 42; Mismatches 99; Indels 18; Gaps 6;
Qy 68 AATTIVASSETKEEDFEQIASLDLNASPLEIMDRALDKFGNDIAIFSGAEDVALLEYA 127
Db 40 AAAAATAAPEGAKATDAAETKELNNKSPLEIMDHATFGDEVAIAIFSGAEDVALLEYA 99
Qy 128 KLGRPRVPSLDTGRNLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFYE 187
Db 100 HLGRKRVPSLDTGRNLNPETYQLFDAVEKHGIRIETFPDAQETMDLVREKGLFSFYE 159
Qy 188 DGHQECRRVKRPLRALKGLRAWITGQRKDQSPGTRSEIPVQVDPAPFEGMDGGIGSL 247
Db 160 DGHQECRRVKRPLRLQLKTLQAWITGQRKDQSPGTRMEVPAVQVDPFEGQSGDGL 219
Qy 248 VKNPNVANVKGHDITNLTMTNVPVNSHAKGYVSGCEPCTRPVLPQGEREGRWMD 307
Db 220 VKYNPLNTMTSAEYVNFIRNVPNSKLHCGYVSGCEPCTRPVLPNQQEREGRWMD 279
Qy 308 AKAKECGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNSPNVNLSTRTGIENLAK 367
Db 280 SAAKECGLHSGNVDSAEAKAE-----APDLWTGGAVALDKATLEKLAG 326
Qy 368 LEDRKEPWLVLVYAPWCPYCOAMBEESYVDLADKLAGSTGMKVGFRADGEOKEPAKSELQ 427
Db 327 -GDRDKDTMVLVYAPWCPFCQAFEPDYEFPVAQE-PGGAGFGVAKYNADAD-REYSSES-LG 382
Qy 428 LGSPTTILFFPKHSSRTTIKYPSEKRDVDSLMAFVNAL 465
Db 383 LKTFPTTIFLPKGS DK-VVKFESDRRTVESMTWMAAAL 419

RESULT 8
H83426
3'-phosphoadenosine-5'-phosphosulfate reductase PA1756 [imported] - Pseudomonas aerugin
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83426
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83426
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <STO>
A;Cross-references: UNIPROT:O05927; UNIPARC:UPI0000128CED; GB:AE004601; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: cydH; PA1756

C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 30.5%; Score 745.5; DB 2; Length 267;
Best Local Similarity 54.8%; Pred. No. 3.2e-50;
Matches 130; Conservative 46; Mismatches 59; Indels 3; Gaps 1;
Qy 83 DFEQIASLDLNASPLEIMDRALDKFGNDIAIFSGAEDVALIEYAKLTGRPFRVFSLDTG 142
Db 29 DLPALASSLADKSPQDILKAAPFHFGEDELWISFGAEDVVLVDMAWLNKRVKFSLDTG 88
Qy 143 RLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFYEDGHQECRRVKVRPL 202
Db 89 RLHPETVRFIDQVREHYGIAIDVLSPDRLPLVKEKGLFSFYRDGHGECGGRKIEPL 148
Qy 203 RRALKGLRAWITGQRKDQSPGTRSEIPVQVDPAPFEGMDGGISLVKWNPNVANVKGHDII 262
Db 149 KRLAGVRAWATGQRDQSPGTRSGVAVLEIDGAFSTPE---KPLYKFNPLSSMTSBEVM 205
Qy 263 NFLRTMTNVPVNSHAKGYVSGCEPCTRPVLPQGEREGRWMDAKAKCEGLHKGNV 320
Db 206 GYIRMLPYNLSHERGYISIGCEPCTRPVLPQGEREGRWMDAEATHKECGLHAGNL 263

RESULT 9
D90470
hypothetical protein cydH [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: D90470
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.W.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90470
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-239 <KUR>
A;Cross-references: UNIPROT:Q97UT3; UNIPARC:UPI000000648C0; GB:AE006641; NID:gl3816282; P
C;Genetics:
A;Gene: cydH
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 21.2%; Score 519.5; DB 2; Length 239;
Best Local Similarity 43.3%; Pred. No. 8.8e-33;
Matches 106; Conservative 48; Mismatches 78; Indels 13; Gaps 7;
Qy 81 EEDFEQIASLDLNASPLEIMDRALDKFGNDIAIFSGAEDVALIE-YAKLTGRPFRVFS 138
Db 4 KDELDLSINKLFEDKEPLEVLKWKIEKYPKIALACSLQAEEDLVILDMLSKYTKRP-RVFI 62
Qy 139 LDTGRNLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFYE--DGHQECRRV 196
Db 63 IDTGRHQESYDLTEEVVKYKYNIDLRIVFPDYKEVEDLVNKYGINFLFKSVELRKACCEV 122
Qy 197 RKVPLRALKGLRAWITGQRKDQSPGTRSEIPVQVDPAPFEGMDGGISLVKWNPNVANV 256
Db 123 RKVPLRALKGLRAWITGQRKDQSPGTRSEIPVQVDPAPFEGMDGGISLVKWNPNVANV 174
Qy 257 KGHDIWNLRTMTNVPVNSHAKGYVSGCEPCTRPVLPQGEREGRWMDAKAKCEGLH 316
Db 175 TWEQWVEYINKNVPVNSHAKGYVSGCEPCTRPVLPQGEREGRWMDAKAKCEGLH 234
Qy 317 KGNVYK 321
Db 235 YREVK 239

RESULT 10
H81905
probable phosphoadenosine phosphosulfate reductase NMA1366 [imported] - Neisseria mening
C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: H81905
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
; Holroyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: H81905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-244 <P>
A;Cross-references: UNIPROT:Q9JUD5; UNIPARC:UPI00000C4BA6; GB:AL162755; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 19.2%; Score 470.5; DB 2; Length 244;
Best Local Similarity 38.0%; Pred. No. 5.7e-29;
Matches 100; Conservative 48; Mismatches 82; Indels 33; Gaps 8;

Qy 58 PPKDSIVPLAATTIVASSETKEEDFEQIASDLNAPLEIMDRALDKFGNDIAAFSG 117
Db 12 PP-----IENGSETALAEKTTETLQRLHRIAGSHRDA-----RFASSLA----- 50

Qy 118 AEDVALIEYAKLTGRPRVPSLDTGRINPETYQLFDAVEKHY-GIRIEYMFPPDAVEYQAL 176
Db 51 AEDWVITDLIAGENLNIGIFLDTGLLHAETLNLDRIVPHMQIKRFPQIREDAHY 110

Qy 177 VRSGKGLSPFYB--DGHQECRCVRKVRPLRALKGLRWITGMVQKQSPGTRSEIPVQVD 234
Db 111 VESKGRFAFYDSVEARRECCIRKTEPLNRAIAGADAWLTQRRREQS-ATRTBLPFAEYD 169

Qy 235 PAPEGMDGGISLVKPNVAVKGDHWNFLRTMNVVPSNLSHAKGYVSIGCEPCTRPVLP 294
Db 170 -AGRGID-----KYNPIFDWSEHDVWYILANNVPYNDLYRQGFPSIGCDGCTRPVKA 221

Qy 295 GQHERGRWWEDAKAKCEGLHK 317
Db 222 GEDIRAGRWWEDKNSKECGLHK 244

RESULT 11
C8111
phosphoadenosine phosphosulfate reductase NMB1193, NMB1155 [imported] - *Neisseria meningitidis*
C;Species: *Neisseria meningitidis*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81115
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81111
A;Molecule type: DNA
A;Residues: 1-246 <TE2>
A;Cross-references: UNIPROT:Q9JRT1; UNIPARC:UPI00000C4DB6; GB:AE002467; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Accession: C81115
A;Molecule type: DNA
A;Residues: 1-246 <TE2>
A;Cross-references: UNIPARC:UPI00000C4DB6; GB:AE002464; GB:AE002098; NID:97226388; PIDN:
A;Experimental source: serogroup B, strain MC58
C;Genetics:
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 18.9%; Score 463; DB 2; Length 246;
Best Local Similarity 37.1%; Pred. No. 2.2e-28;
Matches 101; Conservative 46; Mismatches 93; Indels 32; Gaps 8;

Qy 49 SLVKPVNAEPPRKDSIVPLAATTIVASSETKEEDFEQIASDLNAPLEIMDRALDKFG 108
Db 4 TLFKALWQPHIGS-----GGETALAEKTTETLQRLHRIAGSHRDA-----RFA 48

Qy 109 NDIAIAFSGAEDVALIEYAKLTGRPRVPSLDTGRINPETYQLFDAVEKHY-GIRISYMP 167
Db 49 SSIA-----AEDWVITDLIAGENLNIGIFLDTGLLHAETLNLDRIGRAYPHLRIRKR 103

Qy 168 PDAVEYQALVRSGKGLSPFYB--DGHQECRCVRKVRPLRALKGLRWITGMVQKQSPGTR 225
Db 104 PVREDADRYVESKGRFAFYDSVEARRECCIRKTEPLNRAIAGADAWLTQRRREQS-ATR 162

Qy 226 SEIPVQVDPAFEGMDGGISLVKPNVAVKGDHWNFLRTMNVVPSNLSHAKGYVSIGC 285
Db 163 TELPFAEYDAG-----RGIG-----KYNPIFDWSEHDVWYILANNVPYNDLYRQGFPSIGC 214

Qy 286 EPCTRPVLPQHERGRWWEDAKAKCEGLHK 317
Db 215 DPCTRPVKAEDIRAGRWWEGNRNSKECGLHK 246

RESULT 12
H97458
hypothetical protein AGR_C1497 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
C;Species: *Agrobacterium tumefaciens*
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97458
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Gallinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H97458
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <KUR>
A;Cross-references: UNIPROT:Q8UH67; UNIPARC:UPI0000128CE6; GB:AB007869; PIDN:AAK86625.1;
C;Genetics:
A;Gene: AGR_C1497
A;Map position: circular chromosome
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 16.6%; Score 406; DB 2; Length 260;
Best Local Similarity 35.4%; Pred. No. 6.2e-24;
Matches 92; Conservative 37; Mismatches 117; Indels 14; Gaps 4;

Qy 65 VPLAATTIVASSETKEEDFEQIASDLNAPLEIMDRALDKFGNDIAIAFSGAEDVALI 124
Db 3 IPDVTWTTINSTNASADTASLDATLADLGLSLFV---AGLGGRAVFTTSLGIEDQVIT 59

Qy 125 EYAKLTGRPRVPSLDTGRINPETYQLFDAVEKHYGIRIEYMFPPDAVEYQALVRSGKGLS 184
Db 60 AAGTHTRLPIDVVTLETGRFKEVTDLIDTEERFGIETRRFRPEQDDIDDAYAKYGLNG 119

Qy 185 FYE--DGHQECRCVRKVRPLRALKGLRWITGMVQKQSPGTRSEIPVQVDPAFEGMDG 242
Db 120 FYESVEARHACCHVRKLIPLGKALEGAFFWITGLRRGQS-GNRAATPFAEFDAE----- 172

Qy 243 GIGSLVKPNVAVKGDHWNFLRTMNVVPSNLSHAKGYVSIGCEPCTRPVLPQHERGR 302
Db 173 --RNLIKINALADWDIEQIRAYVAEENIPVNPILHORGYPSIGCEPCTRAIKPGEPEAGR 230

Qy 303 WWEDAKAKCEGLHKGNVQ 322
Db 231 WWENDEKRECGLHVAGAEQ 250

RESULT 13
AB2677
phosphoadenosine phosphosulfate reductase [imported] - *Agrobacterium tumefaciens* (strain
C;Species: *Agrobacterium tumefaciens*
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AB2677

Job time : 41 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:30:19 ; Search time 228 Seconds
(without alignments)
1442.001 Million cell updates/sec

Title: US-10-731-525-8
Perfect score: 2445
Sequence: 1 MALAFTSSISAPTSPTTFSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1951	79.8	470	2	Q8W1A1_SOYBN
2	1883.5	77.0	465	2	Q6V3B0_9ROSI
3	1832.5	74.9	461	2	Q672Q8_LYCERS
4	1829	74.8	464	2	Q39619_CATRO
5	1811.5	74.1	465	1	APR1_ARATH
6	1808	73.9	464	2	Q2ZP22_BRAJU
7	1806.5	73.9	465	2	Q8LA60_ARATH
8	1801	73.7	458	2	Q8LAV2_ARATH
9	1797	73.5	458	1	APR3_ARATH
10	1758.5	71.9	455	2	Q541D4_ARATH
11	1738	71.1	454	1	APR2_ARATH
12	1715	70.1	454	2	Q9ZP23_BRAJU
13	1668.5	68.2	461	2	Q5EUC9_MAIZE
14	1663	68.0	466	2	Q5EUC8_MAIZE
15	1639.5	67.1	475	2	Q6ZAA7_ORYSA
16	1597.5	65.3	459	2	Q9SCA0_LEWMI
17	1494.5	61.1	442	2	Q9SDP3_ALLCE
18	1483	60.7	456	2	Q6PMW2_CERRI
19	1482	60.6	329	2	Q8L6B6_SOLTU
20	1441.5	59.0	465	2	Q8LSD4_PHYPA
21	1313.5	53.7	302	2	Q93Y96_MAIZE
22	1242	50.8	413	2	Q8S2V8_CHLRE
23	1216	49.7	423	2	Q81350_ENTIN
24	840	34.4	192	2	Q5JC58_PEA
25	829	33.9	161	2	Q6DT26_TOBAC
26	785	32.1	249	2	Q6FDX4_AC1AD
27	758.5	31.0	186	2	Q56X39_ARATH
28	745.5	30.5	267	1	CYSH_PSEAE
29	729.5	29.8	244	2	Q4J5Y1_AZOVI
30	728.5	29.8	244	2	Q4KFL0_PSEFS
31	717.5	29.3	244	2	Q6VTB4_9BACT

32	714.5	29.2	244	2	Q88KG2_PSEPK
33	712.5	29.1	244	2	Q883S0_PSESM
34	710.5	29.1	240	2	Q604R9_METCA
35	708.5	29.0	244	2	Q4ZUQ1_PSESV
36	613.5	25.1	171	2	Q9REM5_PSYAN
37	603	24.7	222	2	Q5IWM5_PROWI
38	572	23.4	151	2	Q56YI2_ARATH
39	541	22.1	259	2	Q702F7_9CREN
40	525.5	21.5	239	2	Q96XF0_SULTO
41	519.5	21.2	239	2	Q97UT3_SULSO
42	517.5	21.2	250	2	Q7W9U4_BORPA
43	517.5	21.2	250	2	Q7W9U4_BORPA
44	508	20.8	267	2	Q8ZUM4_PYRAB
45	496	20.3	235	2	Q74CF8_GEOSL

ALIGNMENTS

RESULT 1					
Q8W1A1_SOYBN					
ID	Q8W1A1_SOYBN	PRELIMINARY;	PRT;	470	AA.
AC	Q8W1A1				
DT	01-MAR-2002	(TREMBlrel. 20, Created)			
DT	01-MAR-2002	(TREMBlrel. 20, Last sequence update)			
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)			
DE	Adenosine 5'-phosphosulfate reductase.				
OS	Glycine max (Soybean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
OX	NCBI_TaxID=3847;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Phartiyal P., Krishnan H.B.;				
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF452450; AAL6290.1; -; mRNA.				
DR	GO; GO:0005489; F:electron transporter activity; IEA.				
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.				
DR	GO; GO:0016671; F:oxidoreductase activity, acting on sulfur g. . .; IEA.				
DR	GO; GO:0016740; F:transferase activity; IEA.				
DR	GO; GO:0008652; P:amino acid biosynthesis; IEA.				
DR	GO; GO:0019344; P:cysteine biosynthesis; IEA.				
DR	GO; GO:0006118; P:electron transport; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	GO; GO:0019421; P:sulfate reduction, APS pathway; IEA.				
DR	GO; GO:0006810; P:transport; IEA.				
DR	InterPro; IPR004508; APS_reduc.				
DR	InterPro; IPR002500; PAPS_reduc.				
DR	InterPro; IPR006662; ThioRed.				
DR	InterPro; IPR006663; ThioRedox_dom2.				
DR	Pfam; PF01507; PAPS_reduc; 1.				
DR	Pfam; PF00085; ThioRedoxin; 1.				
DR	PRINTS; PR00421; THIOREDOXIN.				
DR	TIGRFAMs; TIGR00424; APS_reduc; 1.				
SQ	SEQUENCE 470 AA; 51965 MW; 91CSB4PEA67FF064 CRC64;				

Query Match					
Best Local Similarity 80.3%; Pred. No. 7.2e-147;					
Matches 382; Conservative 38; Mismatches 40; Indels 16; Gaps 7;					
Qy	1	MALAF-----TSSISAPTSTF-----PSEPKLPQIGSIRISRPFGAVFNLSQRSL	50		
Db	1	MALAVSTSSSSAAAAAASSFFSRLGSSSDAKAPQIGSFPPERSLVSSVVNVVTVQRSL	60		
Qy	51	VKPVNAEPPRKDSIVPLAATTIVASSETKEDFEQIASOLDNASPLEIMDRALDKFGND	110		
Db	61	VRPLNAEPQRNDSIVPLAA--TIVAPEVEKEEDFEQIAKOLENASPLEIMDRALKEFGND	119		
Qy	111	IATAFSGAEDVALIEYAKLTGRPRVFLSDTGRINPETYQLFDVAKHYGIRIEMFPDA	170		
Db	120	IATAFSGAEDVALIEYAKLTGRPRVFLSDTGRINPETYKFFDAVEKHYGIRIEMFPDA	179		

QY 171 VEVOALVRSKGLRFSYEDGHOECRRVVRPLRRLAKGLRAWITGQRKDSQPGTRSEIPV 230
 DB 180 VEVOALVRTKGLRFSYEDGHOECRRVVRPLRRLAKGLRAWITGQRKDSQPGTRSEIPV 239
 QY 231 VQVDPAPFEGDGGIGSLVKNPVPANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTR 290
 DB 240 VQVDPVFEGLDGGIGSLVKNPVPANVGLDWSFLRTMDVPVNSLHSGQVYSIGCEPCTR 299
 QY 291 PVLPGQHERGRWWEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFN 350
 DB 300 PVLPGQHERGRWWEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFN 356
 QY 351 SPNVNLSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESVVDLADKLKAGSTGMKVG 410
 DB 357 SQDVVLSRSGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESVVDLADKLKAGS-GYKVA 415
 QY 411 KFRADGQKEFAKSELQSGFPTTLFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
 DB 416 KFRADGQKEFAKSELQSGFPTTLFPFKHSSQP-IKYPSEKRDVDSLMAFVNALR 470

RESULT 2
 Q6V3B0_9ROSI PRELIMINARY; PRT; 465 AA.
 ID Q6V3B0_9ROSI PRELIMINARY; PRT; 465 AA.
 AC Q6V3B0;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Adenosine 5' phosphosulfate reductase.
 OS Populus alba x Populus tremula.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
 OX NCBI_TaxID=80863;
 RN [1]
 RP Kopriva S., Hartmann T., Massaro G., Hoenicke P., Rennenberg H.;
 RA "Regulation of sulfate assimilation by nitrogen and sulfur nutrition
 RT in poplar trees."; 2004.
 RL Trees 18:320-326 (2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15020643; DOI=10.1093/ixb/erh094;
 RA Hartmann T., Hoenicke P., Wirtz M., Hell R., Rennenberg H., Kopriva S.;
 RT "Regulation of sulphate assimilation by glutathione in poplars
 RT (Populus tremula x P. alba) of wild type and overexpressing gamma-
 RT glutamylcysteine synthetase in the cytosol."; 2004.
 RL J. Exp. Bot. 55:837-845 (2004).
 DR EMBL: AY353089; AAQ57202.1; -; mRNA.
 DR GO: GO:0005489; F:electron transporter activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0016671; F:oxidoreductase activity, acting on sulfur g. . .; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0019344; P:cysteine biosynthesis; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR GO: GO:0019421; P:sulfate reduction, APS pathway; IEA.
 DR InterPro: IPR004508; APS_reduct.
 DR InterPro: IPR011798; APS_reductase.
 DR InterPro: IPR002500; PAPS_reduct.
 DR InterPro: IPR006662; ThioRedox.
 DR InterPro: IPR006663; ThioRedox_dom2.
 DR Pfam: PF01507; PAPS_reduct; 1.
 DR PRINTS: PR00421; THIOREDOXIN.
 DR TIGRFAMs: TIGR00424; APS_reduct; 1.
 DR TIGRFAMs: TIGR02055; APS_reductase; 1.
 SQ SEQUENCE 465 AA; 51540 MW; 19A9B6A4C934AC21 CRC64;

Query Match 77.0%; Score 1883.5; DB 2; Length 465;
 Best Local Similarity 77.9%; Pred. NO. 1.7e-141;
 Matches 366; Conservative 35; Mismatches 60; Indels 9; Gaps 7;

QY 1 MALAFTSSISAPTSTF--PSSEPKLPQIGSIRISERPIGAVNPNLSQRRSLVKVPVNAEP 58
 DB 1 MALVSSSSISASGFSRSQELKAPQFGSKLLDRET-TFVNVS--SQRCAVKPLNAEP 58
 QY 59 PRKDSIVPLAATTIVASASE--TKBEDFQIADSLDNASPLEIMDRALDKFGNDIAIAFS 116
 DB 59 KRNGSVVPLAATTAAPEIAEKVEVEDEYELAKELVNGSPLEIMDKALEKFGDDIAIAFS 118
 QY 117 GAEDVALIEYAKLTGPRFVPSLDTGRNLNPTETVQLPDAVEKHGIRIERYFPDPDAVEQAL 176
 DB 119 GAEDVALIEYAKLTGPRFVPSLDTGRNLNPTETVQLPDAVEKHGIRIERYFPDPDAVEQAL 178
 QY 177 VRSKGLFSFYEDGHOECRRVVRPLRRLAKGLRAWITGQRKDSQPGTRSEIPVQVDPV 236
 DB 179 VRNKGLESFYEDGHOECRRVVRPLRRLAKGLRAWITGQRKDSQPGTRSEIPVQVDPV 238
 QY 237 FEGDGGIGSLVKNPVPANVKGHDIMNPLRTMNPVNSLHAKGVYSIGCEPCTRPLVPGQ 296
 DB 239 FEGLDGCGAGSLIKWNPMPANVEGDVWVKFLRTMDVPVNSLHSGYISIGCEPCTRPLVPGQ 298
 QY 297 HEREGRWWEEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFNPNVN 356
 DB 299 HEREGRWWEEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFNPNVN 357
 QY 357 LSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESVVDLADKLKAGSTGMKVGKFRADG 416
 DB 358 LSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESVVDLADKLKAGS-GYKVGKFRADG 416
 QY 417 EQKEFAKSELQSGFPTTLFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
 DB 417 DQKFGSKQELQSGFPTTLFPFKHSSRP-IKYPSEKRDVDSLMAFVNALR 465

RESULT 3
 Q672Q8_LYCSES PRELIMINARY; PRT; 461 AA.
 ID Q672Q8_LYCSES PRELIMINARY; PRT; 461 AA.
 AC Q672Q8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Adenyllyl-sulfate reductase precursor.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 OX NCBI_TaxID=4081;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Coker J.S., Vian A., Davies E.;
 RT "Identification, accumulation, and functional prediction of novel
 RT tomato transcripts systemically upregulated after fire damage."; 2005.
 RL Physiol. Plantarum 124:311-322 (2005).
 DR EMBL: AY568717; AAU03359.1; -; mRNA.
 DR GO: GO:0005489; F:electron transporter activity; IEA.
 DR GO: GO:0016671; F:oxidoreductase activity, acting on sulfur g. . .; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR GO: GO:0019421; P:sulfate reduction, APS pathway; IEA.
 DR InterPro: IPR004508; APS_reduct.
 DR InterPro: IPR002500; PAPS_reduct.
 DR InterPro: IPR006662; ThioRedox.
 DR InterPro: IPR006663; ThioRedox_dom2.
 DR Pfam: PF01507; PAPS_reduct; 1.
 DR PRINTS: PR00421; THIOREDOXIN.
 DR TIGRFAMs: TIGR00424; APS_reduct; 1.
 KW Transit peptide.
 FT TRANSIT 1 71 Potential.
 SQ SEQUENCE 461 AA; 51146 MW; F7D6BAD38914A799 CRC64;

Query Match 74.9%; Score 1832.5; DB 2; Length 461;

QY 297 HEREGRWEDAKAKCGLHKGNVKQKEDVNGLSQSHANGDATTVPDIENSPNVN 356
 DB 298 HEREGRWEDAKAKCGLHKGNIKE:-----NTNGNATANVNGTA-TVADIIENSENVN 350
 QY 357 LSRGTIENLAKLEDRKPEMLVLYVAPWPCVQCOAMEESYVDLADKLKAGTGMKVKCPADG 416
 DB 351 LSRQGIENLAKLEDRKPEMLVLYVAPWPCVQCOAMEASFDLADKLKAGS-GVKAKPADG 409
 QY 417 EQKEFAKSEIQLGSPFTILFPFKHSSRTTIKYPSEKRDVDSLMAPVNALR 466
 DB 410 DQXDFAKKEIQLGSPFTILFPFKHSSSRP-IKYPSEKRDVDSLTLNLR 458

RESULT 9
 APR3 ARATH
 ID APR3 ARATH STANDARD; PRT; 458 AA.
 AC P92980: Q38948;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (adenosine 5'-phosphosulfate 3, chloroplast precursor (BC 1.8.4.9)
 DE (adenosine 5'-phosphosulfate 5'-adenylylsulfate sulfotransferase 3)
 DE (APS sulfotransferase 3) (Thioredoxin independent APS reductase 3)
 DE (3'-phosphoadenosine-5'-phosphosulfate reductase homolog 26) (PAPS
 DE reductase homolog 26) (Prt-26).
 GN Name=APR3; Synonyms=PRH26; OrderedLocusNames=At4g21990;
 GN ORFNames=F1N20..90;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=97075173; PubMed=8917599; DOI=10.1073/pnas.93.23.13377;
 RA Gutierrez-Marcos J.F., Roberts M.A., Campbell E.I., Wray J.L.;
 RT "Three members of a novel small gene-family from Arabidopsis thaliana
 able to complement functionally an Escherichia coli mutant defective
 in PAPS reductase activity encode proteins with a thioredoxin-like
 domain and 'APS reductase' activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13377-13382(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=97075174; PubMed=8917600; DOI=10.1073/pnas.93.23.13383;
 RA Setya A., Murillo M., Leustek T.;
 RT "Sulfate reduction in higher plants: molecular evidence for a novel
 5'-adenylylsulfate reductase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13383-13388(1996).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=cv. Columbia;
 RA Chen Y.C., Leustek T.;
 RT "three genomic clones from Arabidopsis thaliana encoding 5'-
 adenylylsulfate reductase.";
 RL (er) Plant Gene Register PGR98-030.
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansong W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delsen M., Fugdomesch P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohseil J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkee W.,

MOOIJMAN P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 Bernseiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McAlay K., Mayes R.,
 RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarsee A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frisman D., Haase D., Lencke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Johnson D.,
 RA Kramer J., Fulton L., Mardis E., Dance M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Grant S., Shondy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesena E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 genome.";
 RL Science 302:842-846(2003).
 CC -I- FUNCTION: Reduces sulfate for Cys biosynthesis. Substrate
 preference is adenosine-5'-phosphosulfate (APS) >> 3'-
 phosphoadenosine-5'-phosphosulfate (PAPS). Uses glutathione or DTT
 as source of protons.
 CC -I- CATALYTIC ACTIVITY: AMP + sulfite + glutathione disulfide =
 CC adenylyl sulfate + 2 glutathione.
 CC -I- ENZYME REGULATION: Stimulated by sodium sulfate > ammonium sulfate
 CC (by similarity).
 CC -I- SUBCELLULAR LOCATION: Chloroplast (Potential).
 CC -I- TISSUE SPECIFICITY: Leaves, roots and stem.
 CC -I- INDUCTION: By sulfate starvation.
 CC -I- DOMAIN: The C-terminal domain may function as glutaredoxin and
 CC mediates the interaction of the enzyme with glutathione (GSH).
 CC Active in GSH-dependent reduction of hydroxyethylidysulfide,
 CC cystine, dehydroascorbate, insulin disulfides and ribonucleotide
 CC reductase (By similarity).
 CC -I- SIMILARITY: Belongs to the APS reductase family.
 CC -I- SIMILARITY: Contains 1 thioredoxin domain.
 CC -I- CAUTION: Ref.2 sequence differs from that shown due to a
 CC frameshift in position 409.

Qy 231 VQVDPAFEGMDGGTSLVKNPVPANVKGHDWFLNFTLWVNVSLHAKGVYSGCEPCTR 290
Db 233 VQVDPVEFGLDGGVSLVKNPVPANVKGHDWFLNFTLWVNVSLHAKGVYSGCEPCTR 292
Qy 291 PVLPGQHERGRWWEWDAKCEKGLHKNVKKQKEEDVNGLSQSHANGDATVTPDIFN 350
Db 293 PVLPGQHERGRWWEWDAKCEKGLHKNVKKQKEEDVNGLSQSHANGDATVTPDIFN 341
Qy 351 SPNVVNLSTRTGIENLAKLEDRKEPWLVLVYAPWPCYQCAWEEVYDLADKLKAGTGMKVG 410
Db 342 SNNVVALSRGCVENLLKLENKKEAWLVLYAPWPCYQCAWEEVYDLADKLKAGTGMKVG 400
Qy 411 KFRADGQKQFPAKSELQGLSFPPTLLFPKHSRTTIKYPSKRDVDSLMAPFVNALR 466
Db 401 KFRADGQKQFPAKSELQGLSFPPTLLFPKHSRTTIKYPSKRDVDSLMAPFVNALR 455

RESULT 11
APR2 ARATH STANDARD; PRT; 454 AA.
AC P92981; O04215; O02554; Q38947;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 5'-adenylsulfate reductase 2, chloroplast precursor (EC 1.8.4.9)
DE (Adenosine 5'-phosphosulfate 5'-adenylsulfate sulfotransferase 2)
DE (APS sulfotransferase 2) (Thioredoxin independent APS reductase 2)
DE (3'-phosphoadenosine-5'-phosphosulfate reductase homolog 43) (PAPS
DE reductase homolog 43) (Prh-43).
GN Name=APR2; Synonym=APSR, PRH43; OrderedLocusNames=At1g62180;
GN ORFNames=F19K23.11;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia;
RX MEDLINE=97075173; PubMed=8917599; DOI=10.1073/pnas.93.23.13377;
RA Gutierrez-Marcos J.F., Roberts M.A., Campbell E.I., Wray J.L.;
RT "Three members of a novel small gene-family from Arabidopsis thaliana
RT able to complement functionally an Escherichia coli mutant defective
RT in PAPS reductase activity encode proteins with a thioredoxin-like
RT domain and 'APS reductase' activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13377-13382 (1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia; TISSUE=Seedling;
RA Min B., Shin K.W., Ye X., Lee S.Y.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia;
RA Gutierrez-Marcos J.F., Campbell E.I., Wray J.L.;
RT "A fourth member of the APS reductase gene family in Arabidopsis
RT thaliana.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia;
RA Chen Y.C., Leustek T.;
RT "Three genomic clones from Arabidopsis thaliana encoding 5'-
RT adenylylsulfate reductase.";
RL (er) Plant Gene Register PGR98-030.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.D., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Dewar K.,

RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Hunter J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Giller J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Sallberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysocka V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820 (2000).
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Havaehizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
RA Khan S., Koseena E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Katsuya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Samu M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846 (2003).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 45-454.
RC STRAIN=cv. Columbia;
RX MEDLINE=97075174; PubMed=8917600; DOI=10.1073/pnas.93.23.13383;
RA Setya A., Murillo M., Leustek T.;
RT "Sulfate reduction in higher plants: molecular evidence for a novel
RT 5'-adenylsulfate reductase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13383-13388 (1996).
CC -1- FUNCTION: Reduces sulfate for Cys biosynthesis. Substrate
CC preference is adenosine-5'-phosphosulfate (APS) >> 3'-
CC phosphoadenosine-5'-phosphosulfate (PAPS). Uses glutathione or DTT
CC as source of protons.
CC -1- CATALYTIC ACTIVITY: AMP + sulfite + glutathione disulfide =
CC adenylyl sulfate + 2 glutathione.
CC -1- ENZYME REGULATION: Stimulated by sodium sulfate > ammonium sulfate
CC (by similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast (Potential).
CC -1- TISSUE SPECIFICITY: Leaves and stem.
CC -1- INDUCTION: By sulfate starvation.
CC -1- DOMAIN: The C-terminal domain may function as glutaredoxin and
CC mediates the interaction of the enzyme with glutathione (GSH).
CC Active in GSH-dependent reduction of hydroxyethyl disulfide,
CC cystine, dehydroascorbate, insulin disulfides and ribonucleotide
CC reductase (by similarity).
CC -1- SIMILARITY: Belongs to the APS reductase family.
CC -1- SIMILARITY: Contains 1 thioredoxin domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; US3866; AAC49563.1; -; mRNA.
CC EMBL; AF023167; AAB80957.1; -; mRNA.
CC EMBL; U96045; AAB57688.1; -; mRNA.
CC EMBL; AF016283; AAC26980.1; -; Genomic DNA.
CC EMBL; AC000375; AAB60764.1; -; Genomic DNA.
CC EMBL; AF360192; AAK25902.1; -; mRNA.

```
DR EMBL; AY040005; AAK64082.1; -; mRNA.
DR EMBL; U56921; AAC26977.1; ALT_INIT; mRNA.
DR PIR; C96648; C96648. APS_reduc.
DR InterPro; IPR004508; APS_reduc.
DR InterPro; IPR002500; PAPS_reduc.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox dom2.
DR InterPro; IPR012336; ThioRedox-like.
DR InterPro; IPR012335; ThioRedoxin_fold.
DR Pfam; PF01507; PAPS_reduc; 1.
DR Pfam; PF00085; ThioRedoxin; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR00424; APS_reduc; 1.
KW Amino-acid biosynthesis; Chloroplast; Cysteine biosynthesis;
KW Multigene family; Oxidoreductase; Redox-active center;
KW Transist peptide.
FT CHAIN 1 66 Chloroplast (Potential).
FT REGION 67 454 5'-adenylylsulfate reductase 2.
FT REGION 341 454 ThioRedoxin.
FT REGION 67 319 Reductase domain.
FT DISULFID 374 377 Redox-active (By similarity).
FT CONFLICT 16 16 S -> T (in Ref. 3).
FT CONFLICT 40 40 S -> N (in Ref. 3).
FT CONFLICT 43 43 S -> A (in Ref. 1).
FT CONFLICT 45 45 R -> G (in Ref. 7).
FT CONFLICT 47 48 YS -> P (in Ref. 1).
FT CONFLICT 57 58 HS -> SHT (in Ref. 3).
FT CONFLICT 65 65 T -> L (in Ref. 3).
FT CONFLICT 79 79 G -> E (in Ref. 3).
FT CONFLICT 107 107 R -> K (in Ref. 3).
FT CONFLICT 111 111 Q -> E (in Ref. 3).
FT CONFLICT 291 291 R -> S (in Ref. 7).
FT CONFLICT 322 322 K -> F (in Ref. 7).
FT CONFLICT 349 349 K -> R (in Ref. 3).
FT CONFLICT 351 351 G -> R (in Ref. 7).
FT CONFLICT 385 385 I -> V (in Ref. 3).
SQ SEQUENCE 454 AA; 50656 MW; 4EB93C1FFC5E4636 CRC64;

Query Match 71.1%; Score 1738; DB 1; Length 454;
Best Local Similarity 73.0%; Pred. No. 6.9e-130;
Matches 346; Conservative 45; Mismatches 55; Indels 28; Gaps 9;

QY 1 MALAFTSSISAPT-STF-----PSSEPKLPIQGISIRISERPIGGAVNFNLSQRRSLVXKPVN 55
DB 1 MALAVTSSSTAISGSSFSRSGPCSDRKALQICSPFLSDLS-----HVSQRRYSLK-AE 52

QY 56 AEP-PRKDSIVPLAATTIVASASET--KEEPEQIASDLNAPLEIMDRALDKFGNDIA 112
DB 54 AESHSRSSEWTRASTLIAPEVEEKGGEVDFEQLAKKLEDPASPLEIMDKALERFGDQIA 113

QY 113 IAFSGAEDVALIEYAKLTGRPRVPSLDTGRNLNPTTYQLFDAVEKHGYIRIEMFPPDAVE 172
DB 114 IAFSGAEDVALIEYARLTGRPRVPSLDTGRNLNPTTYRLFDAVEKHGYIRIEMFPPDAVE 173

QY 173 VOALVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQV 232
DB 174 VOALVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQV 233

QY 233 VDPAFEGMDGGIGSLVKWNPVANVKGHDIMNPLRTMNPVNSLHAKGVYVSGCEPCTRPV 292
DB 234 VDPVFEGLDGGVGLSVKWNPLANVEGADVNFRLTMDVPVNALHAQGVYVSGCEPCTRPV 293

QY 293 LPQGHREGRWWEDAKKEGGLHKGNVKQOKEEDVNGLSQSHANGDATVVDIENSP 352
DB 294 LPQGHREGRWWEDAKKEGGLHKGNT---KEED-----GAADSKPAAVQEIFESN 342

QY 353 NVNLSRRTGIENLAKLEDRKEPFWLVLYAPWCPYQCAEESYVDIADKLAGTGMKVGKF 412
DB 343 NVVALSKGVENLLKLENKEAWLVLYAPWCPYQCAEASYLEAEKLAGK-GVKYAKF 401

QY 413 RADGQKEFAKSEIQLGSEFTTILFPFKHSRPTTIKYPSEKDKDVSLSMAFVNLR 466
DB 402 RADGQKEFAKQELQGLSFTTILFPFKRAPR-AIKYPSEHRDVSLSMFVNLR 454
```

RESULT 12

```
Q9ZP23_BRAJU
ID Q9ZP23_BRAJU PRELIMINARY; PRT; 454 AA.
AC Q9ZP23;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE APS reductase precursor.
GN Name-apsr2;
OS Brassica juncea (Leaf mustard) (Indian mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3707;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Root;
RX MEDLINE=99277600; PubMed=10350097; DOI=10.1023/A:1006169717355;
RA Heiss S., Schaefer H., Haag-Kerwer A., Rausch T.;
RT "Cloning sulfur assimilation genes of Brassica juncea L.: cadmium
RT differentially affects the expression of a putative low-affinity
RT sulfate transporter and isoforms of ATP sulfurylase and APS
RT reductase."
RL Plant Mol. Biol. 39:847-857(1999).
DR EMBL; AJ001207; CA004610.1; -; mRNA.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016671; F:oxidoreductase activity, acting on sulfur g. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0019344; P:cysteine biosynthesis; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0019421; P:sulfate reduction, APS pathway; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004508; APS_reduc.
DR InterPro; IPR002500; PAPS_reduc.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox dom2.
DR Pfam; PF01507; PAPS_reduc; 1.
DR Pfam; PF00085; ThioRedoxin; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR00424; APS_reduc; 1.
KW Transit peptide.
FT TRANSIT 1 47 Potential.
FT CHAIN 48 454 APS reductase.
FT SEQUENCE 454 AA; 50636 MW; 219D246482E455CB CRC64;

Query Match 70.1%; Score 1715; DB 2; Length 454;
Best Local Similarity 72.5%; Pred. No. 4.7e-128;
Matches 343; Conservative 44; Mismatches 60; Indels 26; Gaps 9;

QY 1 MALAFTSSISAPT-STF-----PSSEPKLPIQGISIRISERPIGGAVNFNLSQRRSLVXKPVN 55
DB 1 MALAVTSSSTAISGSSFSRSGPCSDRKALQICSPFLSDLS-----HVSQRRYSLK-AE 52

QY 56 AEP-PRKDSIVPLAATTIVASASETKE--EPEQIASDLNAPLEIMDRALDKFGNDIA 113
DB 53 SPPTRNDLSLTVASTLITPGVEEKEDEVEDFEQLAKKLEASPLEIMDKALQRFSGNIAI 112

QY 114 AFSGAEDVALIEYAKLTGRPRVPSLDTGRNLNPTTYQLFDAVEKHGYIRIEMFPPDAVE 173
DB 113 AFSGAEDVALIEYARLTGRPRVPSLDTGRNLNPTTYRLFDAVEKHGYIRIEMFPPDAVE 172

QY 174 QALVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQV 233
DB 173 QALVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQV 232

QY 234 DPAFEGMDGGIGSLVKWNPVANVKGHDIMNPLRTMNPVNSLHAKGVYVSGCEPCTRPV 293
DB 233 DPAFEGMDGGVGLSVKWNPLANVEGADVNFRLTMDVPVNALHAQGVYVSGCEPCTRPV 292
```


RESULT 15

[illegible]

OC *Oryza sativa* [Japanese cultivated group].
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzeae; *Oryza*.
 OX NCBI TAXID=39947;

Query Match	67.1%	Score	1639.5	DB 2	Length	475			
Best Local Similarity	66.6%	Pred. No.	5.3e-122						
Matches	325	Conservative	53	Mismatches	75	Indels	35	Gaps	11
Qy	1	MALAF	TSSISAPTSTFFSPSEPKLPIQIGSIR--ISERPIGGAVNFNLSORRSLVKPVNAEP	58					
Db	1	MASA-TASIS	SHSVL--RDUKAARIIGNVKQVAAAPAAAGTAAAR-AQAPARAVPPLRAAE	56					
Qy	59	PKDSIVPLA	ATTIVASASETKBE-----DPEQIASDLDLNASPLEIMDRAL	104					
Db	57	PARQ----	PVSASAAAAAAPAAVPAEDAAAAADAPAPAVDYEALAQELQGASPLEIMDRAL	113					
Qy	105	KFGNDIAIAF	SGAEDVALIEYAKLTGPPFRVFSLDTCRLNPEYVQLPDAVEKHGYRIE	164					
Db	114	AMFGSDIAIAF	SGAEDVALIEYAKLTGPPFRVFSLDTCRLNPEYVQLPDKVEKHGYRIE	173					
Qy	165	YMFDPADAVEQ	ALVRASKGLFSYEDGHQBCCRVKRPLRRALKGLRAWITCQRKDQSPGT	224					
Db	174	YMFDPADAVEQ	ALVRAKGLFSYEDGHQBCCRARKVPLRRALRGLRAWITCQRKDQSPGT	233					
Qy	225	RSEIPVQVQDPA	FGMDGGIGSLVKWNPVANVKGHDIWNFLPRTMNVFPVNSLHAKGYVISG	284					
Db	234	RAAIPVQVQDPS	FFGLAGGAGSLVKWNPVANVDGKDVMVTFRLAMDVPVNALHAQGYVISG	293					
Qy	285	CEPCTRPVLPQ	CHEREGRWWEADAKACEGLHKNVKQKKEEDVNGGLSOSH-----ANG	340					
Db	294	CEPCTRPVLPQ	CHEREGRWWEADAKACEGLHKGNI-----DDOGAAAAAAHHRKAGGANG	348					
Qy	341	DATT--VPDIFNS	PNVNLISRTGIENLAKLEDKRPWLVLVYAPMCPYQANMEESYVOLAD	399					
Db	349	NGSAGAPDIP	FSSGGVSLITRAGVENLPLESRAEPWLVLVYAPMCPFCFQOMEASYLELAE	408					

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:37:29 ; Search time 49 Seconds
(without alignments)
786.263 Million cell updates/sec

Title: US-10-731-525-8
Perfect score: 2445
Sequence: 1 MALAFTSSISAPTSTFSSSE.....KYPSEKRDVDSLMFAFNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2445	100.0	466	2	US-09-720-318A-8
2	1923	78.7	472	2	US-09-720-318A-6
3	1811.5	74.1	465	2	US-09-720-318A-12
4	1788.5	73.1	463	2	US-09-720-318A-11
5	1645.5	67.3	463	2	US-09-720-318A-10
6	1295.5	53.0	299	2	US-09-720-318A-4
7	1262.5	51.6	293	2	US-09-720-318A-2
8	811.5	33.2	260	2	US-09-328-352-5189
9	745.5	30.5	269	2	US-09-252-991A-30515
10	540	22.1	240	2	US-09-902-540-15169
11	407.5	16.7	251	2	US-09-540-236-2957
12	347	14.2	236	2	US-10-126-279-28
13	347	14.2	236	2	US-10-286-606-28
14	322	13.2	247	2	US-10-126-279-29
15	322	13.2	247	2	US-10-286-606-29
16	313.5	12.8	254	2	US-10-126-279-30
17	313.5	12.8	254	2	US-10-286-606-30
18	307.5	12.6	248	2	US-09-134-001C-4299
19	305.5	12.5	254	2	US-10-126-279-27
20	305.5	12.5	254	2	US-10-286-606-27
21	256.5	10.5	261	2	US-09-962-357-8
22	250.5	10.2	155	2	US-09-710-279-308
23	206.5	8.4	248	2	US-09-543-681A-7091
24	202	8.3	260	2	US-09-489-039A-9855
25	154	6.3	638	1	US-08-557-122A-38
26	154	6.3	638	2	US-09-262-666-38
27	154	6.3	645	2	US-09-538-092-920

28	150	6.1	496	2	US-09-807-258-27	Sequence 27, Appl
29	145	5.9	511	1	US-08-557-122A-4	Sequence 4, Appli
30	145	5.9	511	2	US-09-262-666-4	Sequence 4, Appli
31	145	5.9	515	1	US-08-557-122A-3	Sequence 3, Appli
32	145	5.9	515	1	US-08-557-122A-34	Sequence 34, Appli
33	145	5.9	515	2	US-09-262-666-3	Sequence 3, Appli
34	145	5.9	515	2	US-09-262-666-34	Sequence 34, Appli
35	145	5.9	3052	1	US-08-557-122A-26	Sequence 26, Appl
36	145	5.9	3052	2	US-09-262-666-26	Sequence 26, Appl
37	140	5.7	489	2	US-09-807-258-10	Sequence 10, Appl
38	139.5	5.7	366	2	US-08-984-919A-11	Sequence 11, Appl
39	139.5	5.7	466	2	US-08-984-919A-33	Sequence 33, Appl
40	139.5	5.7	470	2	US-08-984-919A-55	Sequence 55, Appl
41	139.5	5.7	495	2	US-08-984-919A-47	Sequence 47, Appl
42	139.5	5.7	522	2	US-09-368-588-2	Sequence 2, Appli
43	139	5.7	747	2	US-09-999-833A-459	Sequence 459, App
44	139	5.7	747	2	US-10-020-445A-459	Sequence 459, App
45	138.5	5.7	368	2	US-08-781-420-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-720-318A-8
; Sequence 8, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
; US-09-720-318A-8

Query Match	100.00%;	Score	2445;	DB 2;	Length	466;																																																							
Best Local Similarity	100.0%;	Pred. No.	3.4e-260;	Mismatches	0;	Indels	0;	Gaps	0;																																																				
Matches	466;	Conservative	0;																																																										
Qy	1	MALAF	TSSIS	APTST	FPSS	EPKLP	QIGS	IRIS	ERPI	GGAV	NFNL	SQR	SLVK	VPNA	EP	PP	60																																												
Db	1	MALAF	TSSIS	APTST	FPSS	EPKLP	QIGS	IRIS	ERPI	GGAV	NFNL	SQR	SLVK	VPNA	EP	PP	60																																												
Qy	61	KDSIV	PLAAT	TIV	AS	SETKE	EFQ	IAS	DL	NAS	PLEI	M	D	R	A	L	DKFGNDIA	AT	AF	SG	AE	120																																							
Db	61	KDSIV	PLAAT	TIV	AS	SETKE	EFQ	IAS	DL	NAS	PLEI	M	D	R	A	L	DKFGNDIA	AT	AF	SG	AE	120																																							
Qy	121	VALIE	YAKLT	GR	PF	V	FL	D	T	GL	N	P	E	T	Q	L	P	D	A	V	E	K	H	G	I	R	I	E	T	V	M	P	F	P	A	V	E	Q	A	L	V	R	S	K	180																
Db	121	VALIE	YAKLT	GR	PF	V	FL	D	T	GL	N	P	E	T	Q	L	P	D	A	V	E	K	H	G	I	R	I	E	T	V	M	P	F	P	A	V	E	Q	A	L	V	R	S	K	180																
Qy	181	GLF	S	F	Y	E	D	G	H	Q	E	C	C	R	V	R	K	V	R	P	L	R	R	L	K	G	L	R	A	M	I	T	G	Q	R	K	D	Q	S	P	G	T	R	S	E	I	P	V	Q	V	D	P	A	F	E	G	M	240			
Db	181	GLF	S	F	Y	E	D	G	H	Q	E	C	C	R	V	R	K	V	R	P	L	R	R	L	K	G	L	R	A	M	I	T	G	Q	R	K	D	Q	S	P	G	T	R	S	E	I	P	V	Q	V	D	P	A	F	E	G	M	240			
Qy	241	DG	G	I	S	L	V	K	W	N	P	N	V	A	N	K	G	H	D	I	W	N	F	L	R	T	M	N	P	V	N	S	L	H	A	K	G	V	S	T	G	C	E	P	C	T	R	P	V	L	P	G	O	H	E	R	300				
Db	241	DG	G	I	S	L	V	K	W	N	P	N	V	A	N	K	G	H	D	I	W	N	F	L	R	T	M	N	P	V	N	S	L	H	A	K	G	V	S	T	G	C	E	P	C	T	R	P	V	L	P	G	O	H	E	R	300				
Qy	301	GR	W	W	E	D	A	K	A	K	E	G	L	H	K	G	N	V	K	Q	K	E	D	V	N	G	N	G	L	S	Q	S	H	A	N	G	D	A	T	T	V	P	D	I	F	N	S	P	N	V	N	L	S	R	T	360					
Db	301	GR	W	W	E	D	A	K	A	K	E	G	L	H	K	G	N	V	K	Q	K	E	D	V	N	G	N	G	L	S	Q	S	H	A	N	G	D	A	T	T	V	P	D	I	F	N	S	P	N	V	N	L	S	R	T	360					
Qy	361	G	I	E	N	T	A	K	L	E	D	R	K	E	P	M	L	V	L	V	Y	A	P	W	C	P	Y	Q	A	M	E	E	S	Y	V	D	L	A	D	K	L	A	G	S	T	G	M	K	V	G	K	F	R	A	D	G	E	O	K	E	420

```
|||||
Db 361 GIENLAKLEDRKEPWLVLVYAPWCPYQOAMEESYVDLADKLKLAGSTGMKVGKFRADGEQKE 420
Qy 421 FAKSELQSGFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 421 FAKSELQSGFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466

RESULT 2
US-09-720-318A-6
; Sequence 6, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Glycine max
US-09-720-318A-6

Query Match 78.7%; Score 1923; DB 2; Length 472;
Best Local Similarity 79.4%; Pred. No. 1.2e-202; Indels 16; Gaps 7;
Matches 378; Conservative 37; Mismatches 45;

Qy 1 MALAFTSSIS-----APTSTP-----PSSEPKLPQIGSIIRISERPIGGAVNFNLSQRRSL 50
Db 3 LAVETTSSSAAAAAASSSFSLGSSSDAKAQIGSFRRPQVSSGVNLTQRSS 62
Qy 51 VKPNVAPPRKDSIVPLAATTIVASSETKEDEQATSLDNASPLIEMDRALDKFGND 110
Db 63 VRPLNAEPQRNDSVPLAA-TIVAPEVEKEKEDEFEQLAKOLENSSPLEIMDKALEFGND 121
Qy 111 IAIFSGAEDVALLIYAKLTGRPRVFSLDLGRINLPETYQLFDAVEKHGIRIEMFPDA 170
Db 122 IAIFSGAEDVALLIYAKLTGRPRVFSLDLGRINLPETYKFDVAVEKHGIRIEMFPDA 181
Qy 171 VEQALVRSKGLFSFYEDGHQECRCRVRKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPV 230
Db 182 VEQALVVRTKGLFSFYEDGHQECRCRVRKVRPLRRALKGLKAWITGQRKDQSPGTRSEIPI 241
Qy 231 VQDPAFEGMDGGIGSLVKWNPVANVKGHDINWFLRTMNVPVNSLHAKGYVISIGCEPCTR 290
Db 242 VQDPEVFEGLDGGIGSLVKWNPVANVGLDINWFLRTMNVPVNSLHSGYVISIGCEPCTR 301
Qy 291 PVLPGQHEREGRWWEDAKACEGLHKGNVKQOKEEDVNGLSQSHANGDATTVPDIFN 350
Db 302 PVLPGQHEREGRWWEDAKACEGLHKGNLKQEDAAQLNGMTSQ--GNLSA-TVADIFPI 358
Qy 351 SPNVNLSRGTIENLAKLEDREKPEWLVLVYAPWCPYQOAMEESYVDLADKLKLAGSTGMKVG 410
Db 359 SQNVVLSRSGIENLAKLENKEHVLVLYAPWCRFCQAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGEQKEFAKSELQSGFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGEQKEFAKSELQSGFPTILFPFKHSSQP-IKYPSEKRDVDSLTAFAVNLR 472

RESULT 3
US-09-720-318A-12
; Sequence 12, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
```

```
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-720-318A-12

Query Match 74.1%; Score 1811.5; DB 2; Length 465;
Best Local Similarity 73.9%; Pred. No. 2.3e-190; Indels 19; Gaps 7;
Matches 351; Conservative 48; Mismatches 57;

Qy 1 MALAFTSSISAPTSTPSS-----EPKLPQIGSIIRISERPIGGAVNFNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPKVSGISLRLDRVHVAVPSLNSCKRSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFEQIASDLONASPLIMDRALDKFGNDI 111
Db 61 NAEPKTKDSMIPLAATWVAIEAEVEVVEIEDFEELAKLENASPLIMDKALEKYGNDI 120
Qy 112 IAIFSGAEDVALLIYAKLTGRPRVFSLDLGRINLPETYQLFDAVEKHGIRIEMFPDAV 171
Db 121 IAIFSGAEDVALLIYAKLTGRPRVFSLDLGRINLPETYRFFDAVEKHGIRIEMFPDSV 180
Qy 172 EVQALVRSKGLFSFYEDGHQECRCRVRKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPV 231
Db 181 EVQGLVRSKGLFSFYEDGHQECRCRVRKVRPLRRALKGLKAWITGQRKDQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKWNPVANVKGHDINWFLRTMNVPVNSLHAKGYVISIGCEPCTR 291
Db 241 QVDPVFEGLDGGIGSLVKWNPVANVEGNDVWNLRTMDVPVNTLHAAAGYISIGCEPCTKA 300
Qy 292 VLPQHEREGRWWEDAKACEGLHKGNVKQOKEEDVNGLSQSHANGDATTVPDIFNS 351
Db 301 VLPQHEREGRWWEDAKACEGLHKGNVK-ENSDDAKAVNGESKES-----AVADIFKS 352
Qy 352 PNVNLSRGTIENLAKLEDREKPEWLVLVYAPWCPYQOAMEESYVDLADKLKLAGSTGMKVGK 411
Db 353 ENLVTLRSGIENLAKLENKEPWIIVLYAPWCPFCQAMEASYDELADKLKLAGS-GIKVAK 411
Qy 412 FRADGEQKEFAKSELQSGFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 412 FRADGDQKEFAKQELQSGFPTILVFPKNSRP-IKYPSEKRDVESLTSFLNLVR 465

RESULT 4
US-09-720-318A-11
; Sequence 11, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Catharanthus roseus
US-09-720-318A-11

Query Match 73.1%; Score 1788.5; DB 2; Length 463;
```



```
; Sequence 2, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa = any amino acid
US-09-720-318A-2

Query Match          51.6%; Score 1262.5; DB 2; Length 293;
Best Local Similarity 77.3%; Pred. No. 3.4e-130; Indels 15; Gaps 4;
Matches 235; Conservative 27; Mismatches 27;

QY 167 FPDAVEQALVRSKGLFSFYEDGHECCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRS 226
DB 1 FPDASEVQLVRLKGLFSFYEDGHECCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRA 60

QY 227 EIPVQVDPAFEGWDGIGSLVKWNPVANKGHDINFLRTMNPVNSLHAKGVVSGICE 286
DB 61 SIPVQVDPFEGLDGAGSLVKWNPVANKGKDIWFLRTMDVPVNTLHAQGVVSGICE 120

QY 287 PCTRPVLPGOHEREGRWWEDAKACEGLHKGNVKQKE----EDVNGGLSQSHANGDA 342
DB 121 PCTRPVLPGOHEREGRWWEDAKACEGLHKGNDIKDQAAPRSANGNG-----S 171

QY 343 TTVDPDIFNSPNVNLRTGIENLAKLEDKPELVLVLYAPWCPYCOAMEESYVDLADKLA 402
DB 172 AGAPDIFESPAVSLTGTGIENLLRLNRAEPWLVLYAPWCPYCOAMEASYVELAEKLA 231

QY 403 GSTGMKVKFRADGEQKEPAKSELQLGSFPTLILFPFKHSSRPRTIKYPSEKRDVDSLMAFV 462
DB 232 GS-GVKVAKFRADGEQKPAELQLQSFTPTLLFPQGTARP-IKYPSEKRDVDSLAFV 289

QY 463 NALR 466
DB 290 NSLR 293

RESULT 8
US-09-328-352-5189
; Sequence 5189, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5189
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5189

Query Match          33.2%; Score 811.5; DB 2; Length 260;
Best Local Similarity 57.3%; Pred. No. 1.4e-80; Indels 9; Gaps 3;
Matches 150; Conservative 41; Mismatches 62;

QY 167 FPDAVEQALVRSKGLFSFYEDGHECCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRS 226
DB 1 FPDASEVQLVRLKGLFSFYEDGHECCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRA 60

QY 227 EIPVQVDPAFEGWDGIGSLVKWNPVANKGHDINFLRTMNPVNSLHAKGVVSGICE 286
DB 61 SIPVQVDPFEGLDGAGSLVKWNPVANKGKDIWFLRTMDVPVNTLHAQGVVSGICE 120

QY 287 PCTRPVLPGOHEREGRWWEDAKACEGLHKGNVKQKE----EDVNGGLSQSHANGDA 342
DB 121 PCTRPVLPGOHEREGRWWEDAKACEGLHKGNDIKDQAAPRSANGNG-----S 171

QY 343 TTVDPDIFNSPNVNLRTGIENLAKLEDKPELVLVLYAPWCPYCOAMEESYVDLADKLA 402
DB 172 AGAPDIFESPAVSLTGTGIENLLRLNRAEPWLVLYAPWCPYCOAMEASYVELAEKLA 231

QY 403 GSTGMKVKFRADGEQKEPAKSELQLGSFPTLILFPFKHSSRPRTIKYPSEKRDVDSLMAFV 462
DB 232 GS-GVKVAKFRADGEQKPAELQLQSFTPTLLFPQGTARP-IKYPSEKRDVDSLAFV 289

QY 463 NALR 466
DB 290 NSLR 293

RESULT 9
US-09-252-991A-30515
; Sequence 30515, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30515
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30515

Query Match          30.5%; Score 745.5; DB 2; Length 269;
Best Local Similarity 54.6%; Pred. No. 2.9e-73; Indels 3; Gaps 1;
Matches 130; Conservative 46; Mismatches 59;

QY 83 DFEQIASLDLNASPLEIMDRALDKFGNDIAIAFGAEDVALIEYAKLTGRPRVFSLDTG 142
DB 31 DLPALASSLADKSPQDILKAAFBHFGDELWISFGAEDVVLVDNAWLNRNVKVFSLDTG 90

QY 143 RLNPETYQLFDAVEKHYGIRIETMFPDPAVEQALVRSKGLFSFYEDGHECCRCVRKVRPL 202
DB 91 RLNPETYRFIDQVREHYGIAIDVLSPDRLLEPLVKEKGLFSFYRDGHECCGKIRKIEPL 150

QY 203 RRALKGLRAWITGQRKQSPGTRSEIIPVQVDPAFEGWDGIGSLVKWNPVANKGHDIV 262
DB 151 KRKLAVRAWATGQRDQSPGTRSQAVLEIDGAFSTPE---KPLYKFNPLSSMTSEEV 207

QY 263 NFLRTMNPVNSLHAKGVVSGICEPCTRPVLPGOHEREGRWWEDAKACEGLHKGNV 320
DB 208 GYIRNMLELPYNSLHGRGYISIGCEPCTRPVLPNQHEREGRWWEEATHEKCGLLHAGNL 265

RESULT 10
US-09-902-540-15169
; Sequence 15169, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
```

```
QY 61 KDSIVPLAATTIVASASETKBEDFEIOIASDLDNASPLEIMDRALDKFGNDIAIAFGAED 120
DB 8 RPSLLFGTPTMTVIPTI-----DIVDALAAEAYAKSPREIUELALSQGG-EIALSFGAED 61

QY 121 VALIEYAKLTGRPRVFSLDTGRLNPETYQLFDAVEKHYGIRIETMFPDPAVEQALVRSK 180
DB 62 VVLIDNASRLGKPKPRVFSLDTGLRHPTETQFIETVRKHYNNINIEICFPDAEAQVSMVNEK 121

QY 181 GLPSFYEDGHECCRCVRKVRPLRRALKGLRAWITGQRKQSPGTRSEIIPVQVDPAFEGM 240
DB 122 GLPSFFKDGHECCGIRKVKQPLRKLATLDGMITGQRKQSPGTRTEIIPVQADAGFSGP 181

QY 241 DGGIGSLVKWNPVANKGHDINFLRTMNPVNSLHAKGVVSGICEPCTRPVLPQGHORE 300
DB 182 G---KQLIKYNPLANWSSADVSIYIRNMEIIPYNPLHERGFVSGCEPCTRPVLPNQHERE 238

QY 301 GRWWEDAKACEGLHKGNVKQ 322
DB 239 GRWWEEATQKEGCLHAGNLKK 260
```


; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15169
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15169

Query Match 22.1%; Score 540; DB 2; Length 240;
Best Local Similarity 43.5%; Pred. No. 1.1e-50;
Matches 114; Conservative 43; Mismatches 75; Indels 30; Gaps 8;
QY 58 PPKDSIVPLAATTIVASATKEEDPEQIASLDNASPLEIMDRALDKFGNDIAIAPS- 116
DB 3 PPOFQ-----ASALLAAEAEDVADP-----AQDL-----LAWTER---RFGARAAIASSP 44
QY 117 GAEDVALIEYAKLTGTRFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEYMFDPDAVEQAL 176
DB 45 GVSDMVLIDLARQAPSLRFLTLDTGRLPPTVELMEVVRKRYGVTVETTFPERARVEAL 104
QY 177 VRSGKGLFSFYE--DGHQECRVRKVRPLRRALKGLRAWITGQRKQDSPTGRSEIPVVQVD 234
DB 105 ESTNGYFSPQSLKARACCAIRKVELSRLAQQAQWVTLGRQSV--TRTDVATLEVD 163
QY 235 PAPEGMDGGTSLVKNPVANVKGHDINWFLRTMNPVNSLHAKGYVSGICEPCTRPVLP 294
DB 164 SAH-----GLKLPLATWSRDIWAYVRAKSPYNALHGRGYPISGCAPCTRAVKP 215
QY 295 GQHERGRRWWEDAKKECGH 316
DB 216 YEDERAGRWWESAENRECGH 237

RESULT 11
US-09-540-236-2957
; Sequence 2957, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2957
; LENGTH: 251
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-2957

Query Match 16.7%; Score 407.5; DB 2; Length 251;
Best Local Similarity 40.9%; Pred. No. 4.6e-36;
Matches 88; Conservative 33; Mismatches 77; Indels 17; Gaps 5;
QY 106 KFGNDIAIAPSGAEDVALIEYAKLTGTRFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEY 164
DB 51 KPASSLQ-----AEDMVTIDALANSLSLSTEIFILQTRGLNAETLKLDIVKACYLTINPK 105
QY 165 YMPDPAVEQALVRSGKGLFSFYEDG--HQECRVRKVRPLRRALKGLRAWITGQRKQDSPT 222
DB 106 TYEPFHADVADYVANHGLNAFYSGDLRKLCCFIRKVEPLNRLALVDADAWLTGQRREQSV 165
QY 223 GTSEIPVQVDPAFEGMDGGTSLVKNPVANVKGHDINWFLRTMNPVNSLHAKGYVS 282
DB 166 -TRTELNLAEITDLA-----RGIKYNPFQWQETDVMWAYILTKNIPNELHYGYS 216

QY 283 IGCEPCTRPVLPQOHERGRRWWEDAKAKKECGH 317
DB 217 IGCEPCTMPVKQGEDIRAGRRWWHKKONKECGH 251

RESULT 12
US-10-126-279-28
; Sequence 28, Application US/10126279
; Patent No. 6858213
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012
; CURRENT APPLICATION NUMBER: US/10/126,279
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mycobacterium smegmatis
US-10-126-279-28

Query Match 14.2%; Score 347; DB 2; Length 236;
Best Local Similarity 37.1%; Pred. No. 1.9e-29;
Matches 92; Conservative 34; Mismatches 106; Indels 16; Gaps 7;
QY 70 TTIVASATKEEDPEQIASLDNASPLEIMDRALDKFGNDIAIAPSGAEDVALIEYAKL 129
DB 2 TDVTTSTENELRELAERGAELADASAELLRTDHFHFGNYYVA--SNMQDVLVEMAAK 60
QY 130 TGRPFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEYMFDP--AVEQALVRSGKGLFSFYED 188
DB 61 VRPGVDVLFDITGYHFAETIGTRDAVEYVDVHVNVVTERTVAEQDELGLKGLFA-RDP 119
QY 189 GHQECRVRKVRPLRRALKGLRAWITGQRKQDSPTGRSEIPVQVDPAFEGMDGGTSLV 248
DB 120 G--ECCLRKWVPLTNALKYSAWVTGIRRVAP--TRANAPLISWDNAF-----GLV 168
QY 249 KNPVANVKGHDINWFLRTMNPVNSLHAKGYVSGICEPCTRPVLPQOHERGRRWWEDA 308
DB 169 KINPIAAWTDQMNYIDANGILVNLVYEGYPSIGCAPCTSKPIPCADPRSGR--WAGL 226
QY 309 KAKECGH 316
DB 227 SKTECGH 234

RESULT 13
US-10-286-606-28
; Sequence 28, Application US/10286606
; Patent No. 6863895
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012CIP
; CURRENT APPLICATION NUMBER: US/10/286,606
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279

```
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mycobacterium smegmatis
US-10-286-606-28

Query Match      14.2%; Score 347; DB 2; Length 236;
Best Local Similarity 37.1%; Pred. No. 1.9e-29;
Matches 92; Conservative 34; Mismatches 106; Indels 16; Gaps 7;

Qy 70 TTIVASASEYKEEDFEQIADSLDNASPLEIMDRALDKFGNDIAIFSGAEDVALIEYAKL 129
Db 2 TDVTTSTENELRAERGAELADASABELLRTWDEHFGGNYVVA-SNMQDAVLVEMAAK 60

Qy 130 TGRPFYFSLDTGRLNPTETQYLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYED 188
Db 61 VRPGVDVFLDTGTHFAETITGTRDAVEAVDVHVVTPTVAEQDELLGKNLFA-RDP 119

Qy 189 GHQCCRVKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPVVQVDPAPFGMDGGIGSLV 248
Db 120 G--ECCRLRKVPPLTNALKGYSAMVTGIRRVKAP--TRANAPLISWDNAF-----GLV 168

Qy 249 KWNPNVNVKHDHWNFLRTMNVNPNLSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEA 308
Db 169 KINPIAAWTDEDMQYIDANGILYNPLIYEGYSGICAPCTSKPIPGADPRSGR--WAGL 226

Qy 309 KAKECGLH 316
Db 227 SKTECGLH 234

RESULT 14
US-10-126-279-29
; Sequence 29, Application US/10126279
; Patent No. 6858213
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012
; CURRENT APPLICATION NUMBER: US/10/126,279
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-126-279-29

Query Match      13.2%; Score 322; DB 2; Length 247;
Best Local Similarity 35.8%; Pred. No. 1.2e-26;
Matches 86; Conservative 36; Mismatches 92; Indels 26; Gaps 8;

Qy 88 ASDLDNASPLEIMDRALDKFGN-----DIAIAFSGAEDVALIEYAKLTGRPPRVF 137
Db 21 AAELGASASDVLRWTDFTGGVNGPRGMATCNVVA-SSMQEAVLIDLAARKVRPGVPV 79

Qy 138 SLDTGRNLNPTETQYLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYEDHQECRV 196
Db 80 FLDTGYHFAETITGTRDAIESYDILRVNLVTPHESVAEQDKLLGKDLFA-RDPG--ECCRL 136

Qy 197 RKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPVVQVDPAPFGMDGGIGSLVKNPNVANV 256
Db 137 RKVAPLKGKTLRGYSAMVTGLRRSEA-ATRAPAPVIGFDEGFK-----LVKVPNPMATW 187

Qy 257 KGHDIWNFLRTMNVNPNLSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEAKAKECGLH 316
Db 188 TDEDVQNYIDEHNVNPLIYEGYSGICAPCTAKPLAGADPRSGR--WQGLAKTECGLH 245

Search completed: February 18, 2006, 14:38:50
Job time : 50 secs

; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mycobacterium smegmatis
US-10-286-606-29

Query Match      14.2%; Score 347; DB 2; Length 236;
Best Local Similarity 37.1%; Pred. No. 1.9e-29;
Matches 92; Conservative 34; Mismatches 106; Indels 16; Gaps 7;

Qy 70 TTIVASASEYKEEDFEQIADSLDNASPLEIMDRALDKFGNDIAIFSGAEDVALIEYAKL 129
Db 2 TDVTTSTENELRAERGAELADASABELLRTWDEHFGGNYVVA-SNMQDAVLVEMAAK 60

Qy 130 TGRPFYFSLDTGRLNPTETQYLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYED 188
Db 61 VRPGVDVFLDTGTHFAETITGTRDAVEAVDVHVVTPTVAEQDELLGKNLFA-RDP 119

Qy 189 GHQCCRVKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPVVQVDPAPFGMDGGIGSLV 248
Db 120 G--ECCRLRKVPPLTNALKGYSAMVTGIRRVKAP--TRANAPLISWDNAF-----GLV 168

Qy 249 KWNPNVNVKHDHWNFLRTMNVNPNLSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEA 308
Db 169 KINPIAAWTDEDMQYIDANGILYNPLIYEGYSGICAPCTSKPIPGADPRSGR--WAGL 226

Qy 309 KAKECGLH 316
Db 227 SKTECGLH 234

RESULT 15
US-10-286-606-29
; Sequence 29, Application US/10286606
; Patent No. 6863895
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012CIP
; CURRENT APPLICATION NUMBER: US/10/286,606
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-286-606-29

Query Match      13.2%; Score 322; DB 2; Length 247;
Best Local Similarity 35.8%; Pred. No. 1.2e-26;
Matches 86; Conservative 36; Mismatches 92; Indels 26; Gaps 8;

Qy 88 ASDLDNASPLEIMDRALDKFGN-----DIAIAFSGAEDVALIEYAKLTGRPPRVF 137
Db 21 AAELGASASDVLRWTDFTGGVNGPRGMATCNVVA-SSMQEAVLIDLAARKVRPGVPV 79

Qy 138 SLDTGRNLNPTETQYLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYEDHQECRV 196
Db 80 FLDTGYHFAETITGTRDAIESYDILRVNLVTPHESVAEQDKLLGKDLFA-RDPG--ECCRL 136

Qy 197 RKVRPLRRALKGLRAWITGQRKDQSPGTRSEIPVVQVDPAPFGMDGGIGSLVKNPNVANV 256
Db 137 RKVAPLKGKTLRGYSAMVTGLRRSEA-ATRAPAPVIGFDEGFK-----LVKVPNPMATW 187

Qy 257 KGHDIWNFLRTMNVNPNLSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEAKAKECGLH 316
Db 188 TDEDVQNYIDEHNVNPLIYEGYSGICAPCTAKPLAGADPRSGR--WQGLAKTECGLH 245

Search completed: February 18, 2006, 14:38:50
Job time : 50 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:48:35 ; Search time 168 Seconds
(without alignments)
1158.979 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPTSTPSSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2445	100.0	466	4	US-10-731-525-8
2	2445	100.0	466	4	US-10-762-049-8
3	1951	79.8	470	4	US-10-424-599-277584
4	1951	79.8	475	4	US-10-425-114-43990
5	1923	78.7	472	4	US-10-731-525-6
6	1923	78.7	472	4	US-10-762-049-6
7	1811.5	74.1	465	4	US-10-731-525-12
8	1811.5	74.1	465	4	US-10-762-049-12
9	1788.5	73.1	463	4	US-10-731-525-11
10	1788.5	73.1	463	4	US-10-762-049-11
11	1645.5	67.3	463	4	US-10-731-525-10
12	1645.5	67.3	463	4	US-10-762-049-10
13	1639.5	67.1	475	4	US-10-437-363-126234
14	1622.5	66.4	473	5	US-10-739-930-8328
15	1597.5	65.3	459	4	US-10-169-667A-2
16	1509.5	61.7	386	4	US-10-425-115-267885
17	1444.5	59.1	407	4	US-10-425-115-267882
18	1422.5	58.2	339	4	US-10-435-115-267886
19	1295.5	53.0	299	4	US-10-731-525-4
20	1295.5	53.0	299	4	US-10-762-049-4
21	1262.5	51.6	293	4	US-10-731-525-2
22	1262.5	51.6	293	4	US-10-762-049-2
23	1241.5	50.8	323	4	US-10-425-114-56536
24	1110.5	45.4	227	4	US-10-424-599-278242
25	1045.5	42.8	251	4	US-10-425-115-267884
26	999.5	40.9	252	4	US-10-425-115-267887
27	914	37.4	182	4	US-10-425-115-331871

28	871.5	35.6	213	4	US-10-767-701-42399	Sequence 42399, A
29	754	30.8	147	4	US-10-424-599-225761	Sequence 225761, A
30	731.5	29.9	235	4	US-10-369-493-13795	Sequence 13795, A
31	606	24.8	248	4	US-10-425-115-197638	Sequence 197638, A
32	539	22.0	230	4	US-10-369-493-19207	Sequence 19207, A
33	537	22.0	109	4	US-10-767-701-49099	Sequence 49099, A
34	454.5	18.6	184	4	US-10-424-599-212812	Sequence 212812, A
35	449.5	18.4	225	4	US-10-369-493-19679	Sequence 19679, A
36	410	16.8	158	4	US-10-425-115-140785	Sequence 140785, A
37	393.5	16.1	219	4	US-10-369-493-9785	Sequence 9785, A
38	370.5	15.2	220	4	US-10-369-493-11025	Sequence 11025, A
39	347	14.2	236	4	US-10-126-279-28	Sequence 28, Appl
40	347	14.2	236	4	US-10-286-606-28	Sequence 28, Appl
41	347	14.2	236	5	US-10-891-383-28	Sequence 28, Appl
42	325	13.3	196	4	US-10-369-493-17686	Sequence 17686, A
43	322	13.2	247	4	US-10-126-279-29	Sequence 29, Appl
44	322	13.2	247	4	US-10-286-606-29	Sequence 29, Appl
45	322	13.2	247	5	US-10-891-383-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1

US-10-731-525-8

; Sequence 8, Application US/10731525

; Publication No. US20040121440A1

; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Allen, Stephen M.

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-C

; CURRENT APPLICATION NUMBER: US/10/731,525

; CURRENT FILING DATE: 2003-12-09

; PRIOR APPLICATION NUMBER: US/09/720,318A

; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 466

; TYPE: PRT

; ORGANISM: Glycine max

US-10-731-525-8

Query Match	100.0%;	Score 2445;	DB 4;	Length 466;
Best Local Similarity	100.0%;	Pred. No. 1.1e-222;		
Matches	466;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	MALAFTSSISAPTSTPSSSEPKLPQIGSIRISRPITGGAVNFNLSQRSLVKPVNASPPR	60	
Db	1	MALAFTSSISAPTSTPSSSEPKLPQIGSIRISRPITGGAVNFNLSQRSLVKPVNASPPR	60	
Qy	61	KDSIVPLAATTIVASASSETKEEDFEQIASDLNASPLEIMDRALDKFGNDIAFAFSAED	120	
Db	61	KDSIVPLAATTIVASASSETKEEDFEQIASDLNASPLEIMDRALDKFGNDIAFAFSAED	120	
Qy	121	VALIEYAKLTGRPFVFLDTGRNLNPTETYLQFAVEKHGYIRIEMPPDPAVEQVALVRSK	180	
Db	121	VALIEYAKLTGRPFVFLDTGRNLNPTETYLQFAVEKHGYIRIEMPPDPAVEQVALVRSK	180	
Qy	181	GLFSFYEDGHQECRCRVRPLRRALKGLRAWITGQRKQOSPGRSRIPVVQVDPAPFEGM	240	
Db	181	GLFSFYEDGHQECRCRVRPLRRALKGLRAWITGQRKQOSPGRSRIPVVQVDPAPFEGM	240	
Qy	241	DGGIGSLVKWNPVANKGHDDIWNFLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE	300	
Db	241	DGGIGSLVKWNPVANKGHDDIWNFLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE	300	
Qy	301	GRWWEDAKAKECGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNSPNVNLNLSRT	360	
Db	301	GRWWEDAKAKECGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNSPNVNLNLSRT	360	

Qy 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420
Db 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420
Qy 421 FAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 466
Db 421 FAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 466
RESULT 2
US-10-762-049-8
; Sequence 8, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
US-10-762-049-8

Query Match 100.0%; Score 2445; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.1e-222;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALAFTSSISAPTSTFFSSSEPKLPQIGSIRISERPIGGAVNFNLQSRRSLVKPNABPPR 60
Db 1 MALAFTSSISAPTSTFFSSSEPKLPQIGSIRISERPIGGAVNFNLQSRRSLVKPNABPPR 60
Qy 61 KDSIVPLAATTIVASATSEKEDFEQIASDLDNASPLEIMDRALDKFGNDIAIAFSAED 120
Db 61 KDSIVPLAATTIVASATSEKEDFEQIASDLDNASPLEIMDRALDKFGNDIAIAFSAED 120
Qy 121 VALIEYAKLTGRPFVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDAVEVQALVRSK 180
Db 121 VALIEYAKLTGRPFVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDAVEVQALVRSK 180
Qy 181 GLFSFYEDGHOECRCVRKVRPLRALKGLRAWITGQRKQSPGTRSEIPVQVDPAPFGM 240
Db 181 GLFSFYEDGHOECRCVRKVRPLRALKGLRAWITGQRKQSPGTRSEIPVQVDPAPFGM 240
Qy 241 DGGIGSLVKMNPVANVKGHD1WNFLRTMNPVNSLHAKGVYSIGCEPCTRVLVPGQHERE 300
Db 241 DGGIGSLVKMNPVANVKGHD1WNFLRTMNPVNSLHAKGVYSIGCEPCTRVLVPGQHERE 300
Qy 301 GRWWEDAKACEGLHKGNVKQKEEDVNGNGLSQSHANGDATTVPDIFNSPNVNLNLSRT 360
Db 301 GRWWEDAKACEGLHKGNVKQKEEDVNGNGLSQSHANGDATTVPDIFNSPNVNLNLSRT 360
Qy 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420
Db 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420
Qy 421 FAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 466
Db 421 FAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 466

RESULT 3
US-10-424-599-277584
; Sequence 277584, Application US/10424599

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 277584
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92682C.1.pep
US-10-424-599-277584
Query Match 79.8%; Score 1951; DB 4; Length 470;
Best Local Similarity 80.3%; Pred. No. 8.3e-176;
Matches 382; Conservative 38; Mismatches 40; Indels 16; Gaps 7;
Qy 1 MALAF-----TSSISAPTSTF-----PSSEPKLPQIGSIRISERPIGGAVNFNLQSRRSL 50
Db 1 MALAVTTSSSSAAAASSFFSRLLGSSSDAKAPQIGSFPPERSLVSSVVVNTQRRSL 60
Qy 51 VKPVNAEPPRKDSIVPLAATTIVASATSEKEDFEQIASDLDNASPLEIMDRALDKFGND 110
Db 61 VRPLNAEPQRNDSIVPLAA-TIVAPEVEKEEDFEQIAKOLENASPLEIMDRALDKFGND 119
Qy 111 IATAFSGAEDVALIEYAKLTGRPFVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDA 170
Db 120 IATAFSGAEDVALIEYAKLTGRPFVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDA 179
Qy 171 VEQVALVRSKGLFSFYEDGHOECRCVRKVRPLRALKGLRAWITGQRKQSPGTRSEIPV 230
Db 180 VEQVALVRSKGLFSFYEDGHOECRCVRKVRPLRALKGLRAWITGQRKQSPGTRSEIPV 239
Qy 231 VOVDPAFEGDGGIGSLVKMNPVANVKGHD1WNFLRTMNPVNSLHAKGVYSIGCEPCTR 290
Db 240 VOVDPAFEGDGGIGSLVKMNPVANVKGHD1WFSFLRTMDVFNLSHQYVSGICEPCTR 299
Qy 291 PVLPGQHERGRWWEDAKACEGLHKGNVKQKEEDVNGNGLSQSHANGDATTVPDIFN 350
Db 300 PVLPGQHERGRWWEDAKACEGLHKGNVKQKEEDVNGNGLSQSHANGDATTVPDIFN 356
Qy 351 SPNVNLSRTGIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 410
Db 357 SQDVLSRSRGIENLAKLEDKKEPWLVLVYAPWCRFCQAMEESYVDLAEKLAGS-GYKVA 415
Qy 411 KFRADGQKEFAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 466
Db 416 KFRADGQKEFAKSELQSGSPPTTLFFPKHSSRRPTIKYPSKRDVDSLMAFVNALR 470

RESULT 4
US-10-425-114-43990
; Sequence 43990, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28


```
Db 182 VEQALVTRTKGLSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRGEIPI 241
Qy 231 VQVDPAPFEGMDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 290
Db 242 VQVDPVPEGLDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 301
Qy 291 PVLPGQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIFN 350
Db 302 PVLPGQHERGRWWEDAKAKCEGLHKGNGLKQEDAAQLNGNGTSQ--GNLSA-TVADIFI 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVLSLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGQKFAKSELQSGFPTTLFPFKSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGQKFAKSELQSGFPTTLFPFKSSQ- IKYPSEKRDVDSLTAFAVNALR 472

RESULT 7
US-10-731-525-12
; Sequence 12, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10731.525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-731-525-12

Query Match 74.1%; Score 1811.5; DB 4; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.4e-162;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPTIGGVNPNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPKVQIGSLRLDLDRVHVAPVSLNLSKRSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 111
Db 61 NAEPTKDSIMPLAATVVAETAEVEVEVEEDFELAKLENASPLIMDKALEKYGNDI 120
Qy 112 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLNMPETYQLFDVAKHYGIRIEMFDDAV 171
Db 121 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLNMPETYQLFDVAKHYGIRIEMFDDAV 180
Qy 172 EVQALVRSKGLFSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 231
Db 181 EVQGLVRSKGLFSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 291
Db 241 QVDPVPEGLDGGVGLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 300
Qy 292 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIFNS 351
Db 301 VLPQHERGRWWEDAKAKCEGLHKGNGVK--ENSDDAKNGESK-----AVADIFKS 352
Qy 352 PNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 411
Db 359 SQNVLSLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGQKFAKSELQSGFPTTLFPFKSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGQKFAKSELQSGFPTTLFPFKSSQ- IKYPSEKRDVDSLTAFAVNALR 472
```

```
Db 182 VEQALVTRTKGLSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRGEIPI 241
Qy 231 VQVDPAPFEGMDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 290
Db 242 VQVDPVPEGLDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 301
Qy 291 PVLPGQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIFN 350
Db 302 PVLPGQHERGRWWEDAKAKCEGLHKGNGLKQEDAAQLNGNGTSQ--GNLSA-TVADIFI 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVLSLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGQKFAKSELQSGFPTTLFPFKSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGQKFAKSELQSGFPTTLFPFKSSQ- IKYPSEKRDVDSLTAFAVNALR 472

RESULT 7
US-10-731-525-12
; Sequence 12, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10731.525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-731-525-12

Query Match 74.1%; Score 1811.5; DB 4; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.4e-162;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPTIGGVNPNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPKVQIGSLRLDLDRVHVAPVSLNLSKRSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 111
Db 61 NAEPTKDSIMPLAATVVAETAEVEVEVEEDFELAKLENASPLIMDKALEKYGNDI 120
Qy 112 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLNMPETYQLFDVAKHYGIRIEMFDDAV 171
Db 121 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLNMPETYQLFDVAKHYGIRIEMFDDAV 180
Qy 172 EVQALVRSKGLFSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 231
Db 181 EVQGLVRSKGLFSFYEDGHCCKVRKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 291
Db 241 QVDPVPEGLDGGVGLVKWNPVANKVGHDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 300
Qy 292 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIFNS 351
Db 301 VLPQHERGRWWEDAKAKCEGLHKGNGVK--ENSDDAKNGESK-----AVADIFKS 352
Qy 352 PNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 411
Db 359 SQNVLSLSRGTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAEKLARS-GVKVA 417
```

```

; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Catharanthus roseus
US-10-731-525-11

Query Match      73.1%; Score 1788.5; DB 4; Length 463;
Best Local Similarity 73.4%; Pred. No. 2.1e-160;
Matches 348; Conservative 49; Mismatches 58; Indels 19; Gaps 6;

QY 1 MALAFTSSIS---APTSTPPSPSEPKLPQIGSIRISERPIGGAVNPNLSQRRSLVKPVNAE 57
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MALAFTSSIAIHGSLSSSEFQTKAAQAQFGSQPLDRPHTTISPSVNVSRRLAVKPINAE 60
QY 58 PPKOSIVPLAATTIVASSETKE--EDFEQIASDLDNASPLEIMDRALDKFGNDIAIAFS 116
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 PKRNESIVPSAATTVAPEVEEKVDVEYKLADELQNASPLEIMDKSLAKFGNDIAIAFS 120
QY 117 GAEDVALIEYAKLTGRPRVPSLDTGRNLNPETYQLFDAVEKHGIRIYMPFPDAVEQAL 176
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 GAEDVALIEYAKLTGRPRVPSLDTGRNLNPETYKFFDVEKQYGIHIYMPFPDAVEQAL 180
QY 177 VRSKGLFSYEDHGECCRVKVRPLRALKGLRAWITGQRKQDSGPTSRSEIPVQVDDPA 236
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 VRSKGLFSYEDHGECCRVKVRPLRALKGLRAWITGQRKQDSGPTSRSEIPVQVDDPV 240
QY 237 FEGMDGGISLVKPNVAVNKGHDINWFLRTMNVPVNSLHAKGYVSGICEPTTRPVLPGQ 296
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 FEGMDGGVSLVKPNVAVNVEGDIWNFLRAMDVPVNTLHSGQYVSGICEPTTRPVLPGQ 300
QY 297 HEREGRWEDAKAKCEGLHKGNVKQK----EEDVNGLSQSHANGDATVDPIDFNSP 352
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 HEREGRWEDAKAKCEGLHKGDIKEGTLIIWDGAVNGNG-----SDTIADIFDTN 351
QY 353 NVNLSRGTGIEMLAKLEDRKEPWLVLVAPWCPYQAMSESVVDLADKLKAGSTGMKVGKF 412
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 NTSLSRPGIENLLKLEERREAWLVLYAPWCRFCQAMEGSLYLAELKLAGS-GVYKGF 410
QY 413 RADGEQKEPAKSELQGSFPTTLFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 411 KADGQKAPAQQLQNLSSPTILFPFKHSSKP-IKYPSEKRDVDSLMAFVNALR 463

RESULT 10
US-10-762-049-11
; Sequence 11, Application US/10762049
; Publication No. US2004013942A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Catharanthus roseus
US-10-731-525-11

Query Match      67.3%; Score 1645.5; DB 4; Length 463;
Best Local Similarity 68.8%; Pred. NO. 7.9e-147;
Matches 327; Conservative 49; Mismatches 78; Indels 21; Gaps 8;

QY 1 MALAFTSSISAPTSTPPSPSEPKLPQIGSIR--ISERPIGGAVNPNLSQRRSLVKPVNAEP 58
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MASA-TASIS--SHSIALRDLKAAARIGVAVQVAVPAGLPATAPKQGRARAVRPLCAAE 57
QY 59 PPKOSIVPLAATTIVA-----SASETKEDDFEOLASDLNAPSLEIMDRALDKFGNDIAIA 114

```


THIS PAGE LEFT BLANK

Result No.	Query			DB	ID	Description
	Score	Match	%			
1	250.5	10.2	155	6	US-10-793-626-308	Sequence 308, App
2	154	6.3	645	6	US-10-821-234-1409	Sequence 1409, App
3	139	5.7	747	6	US-10-131-826A-426	Sequence 426, App
4	127.5	5.2	505	6	US-10-821-234-1287	Sequence 1287, App
5	117	4.8	440	6	US-10-821-234-1282	Sequence 1282, App
6	110.5	4.5	324	6	US-10-878-556A-114	Sequence 114, App
7	107.5	4.4	269	7	US-11-000-463-344	Sequence 344, App
8	105.5	4.3	627	6	US-10-467-657-5432	Sequence 5432, App
9	104	4.3	176	6	US-10-821-234-860	Sequence 860, App
10	101.5	4.2	414	6	US-10-878-556A-1	Sequence 1, Appl
11	99	4.0	875	7	US-11-024-959-352	Sequence 352, App
12	91.5	3.7	183	6	US-10-467-962B-6	Sequence 6, Appl
13	91.5	3.7	844	6	US-10-763-712A-48	Sequence 48, Appl
14	90	3.7	310	7	US-11-156-084-227	Sequence 227, App
15	89.5	3.7	760	6	US-10-858-730-76	Sequence 76, Appl
16	88.5	3.6	732	6	US-10-954-468-51	Sequence 51, Appl
17	87.5	3.6	740	6	US-10-821-234-1464	Sequence 1464, App
18	85.5	3.5	408	6	US-10-763-712A-67	Sequence 67, Appl
19	85.5	3.5	456	6	US-10-763-712A-66	Sequence 66, Appl
20	85.5	3.5	610	7	US-11-069-642-22	Sequence 22, Appl
21	85	3.5	280	6	US-10-967-457-75	Sequence 75, Appl
22	85	3.5	440	6	US-10-525-710-52	Sequence 52, Appl
23	85	3.5	1102	7	US-11-098-686-10951	Sequence 10951, A
24	84.5	3.5	1735	7	US-11-040-472-13	Sequence 13, Appl
25	84	3.4	419	7	US-11-156-084-122	Sequence 122, App

```
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1409
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1409

Query Match      6.3%; Score 154; DB 6; Length 645;
Best Local Similarity 19.6%; Pred. No. 9.9e-06;
Matches 114; Conservative 82; Mismatches 192; Indels 194; Gaps 28;

QY 7 SSIAPTSTFSEPKLPQIGSIRISRPICGAVNLSQRRSLVKPVNAEPKRSIVP 66
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 AKIDATSASVLSRFDVSGYPTTKLKK--GQAVDVEGSRTOEIVAKREVSPQDWTPP 175
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 LAATTIVASASETKEEDFEQIASLDL-----NASPLEIMDRALDKFGND 110
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
176 PEVLVLT-----KENFDEVNDADILLVEFVAPWCCHCKLAPEYKAAKELSKRSPP 229
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
111 IAIA-FSGAEVDALIEYAKLTGR-PRVFLSDTGRNLNPETYQLFDVAKHYGIRIEYMF- 167
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 IPLAKVDATAETDLAKRFDVSGYPTLKIFR--NGR--PYDY--NGPREKYGI--VDYME 281
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
168 ---PDAVEQALVRSKGLFSFYEDGHQECRCVRKRLRALKGLRAWITGQRKDSPT 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 QSGPPSKSEILT---KQVSEFLKDGDD-----VILIGVFKGES--- 316
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 RSEIPVVQVDPAPRGMDGI-----GSLV-----KW 250
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 -----DPAYQOQDAANLBEDYKPHHTSTETAKFLKVSQGLVMQPEKFSKY 367
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
251 NP---VANVRGHD-----IMNFLRTMNPV-----NSLHAKGYVSIIGCEPCTRPVLPGQHE 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
368 EPRSHMDVQSGTQDSAIKDFVLKVALPLVGHKRVSNDAKRYT-----RRPLVVVYYS 420
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 -----REGRWV-----ED--AKAKECGLHKNVKKOKEEDV 328
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 VDFSFDYRAATQFWRSKVLEAKDFPEYTFATIADEEDYAGEVKDLGL-----SESGEDV 474
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
329 N-----GNGLSQSHANGDATT-----VPIFNSPNVNLSRT 360
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
475 NAAILDESCKKFAPEPEFSDTLREFTAFKGLKLPVKSQPVKNNKGPVKVVVGKT 534
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 GIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVGKFRADGEQKE 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
535 -FDSI--VMDPKDVLIEFYAPWCGHCKQLEPVPVNSLAKYKQKGLVIAKM--DATAND 589
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 FAKSELQIGSPPTILFPFKSSRRTIKYVPSEKRDVDSIMAFV 462
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
590 VPSDRYKVEGFPPTIYFAPSGDKKNPVKPEGGDRDLHLKFI 631

RESULT 3
US-10-131-826A-426
; Sequence 426, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
```

```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 426
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-426

Query Match      5.7%; Score 139; DB 6; Length 747;
Best Local Similarity 21.8%; Pred. No. 0.00027;
Matches 105; Conservative 70; Mismatches 214; Indels 92; Gaps 23;

QY 27 GSIRISRPICGAVNLSQRRSLVKPVNAEP-----RKDSIVPLAA 69
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 GLLRI-----GAVNCGDDRMLCRMKGVNSYPSLIFPRSGMAPVKYHGDRLSKESLVSPAM 230
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70 TTIVASASETKEEDFEQIASLDLNASPLEIMDRALDKFGNDIAAFSGAEVDALIEYAKL 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 QHVRSTVTELTWGNF---VNSIQTFAGA-----GMLITPCSGGDCITTSQTRL 279
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
130 TGRPPFRVFLSDTGRNLNPETYQLFDVAKHYGIRIEYMFDAVEVQA-----LVRSKGLF 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
280 SGMFL-LSLDA-----KEYLVEVHNLDPFELLSANTLEDRLAHRWLL 323
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 SFYEDGHQECRCVRKRVPLRALKGLRAWITGQRKDSPTGRSIPVQVDDPA-FEGM-- 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 -FFHFGKNVNSNDPELKKLTKLLKNDHIQVGRFDCSSAPDICSNLVYFQPSLAVFKGGQT 382
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 ---DGGIGSLVKNPVA---NVKGDHINPLRTMNPVNSLH--AKGYVSIIGCEPCTRP 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 KEYEIHGHKKILYDILAFAKESVNSH--VTTLGPQNFANDKEPWLVDFFAPWCPPC-RA 439
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
292 VLPQHEREGRWWEEDAK--AKECGLHKG-----NVKQOKEEDV--NGNGLSQSHANGDAT 343
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
440 LLP-ELRRASNLLYGQLKFGTLDTVHEGLCNMNYIOAYPTTVVFNQSNIIHEYEGHSAE 498
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
344 T----VPDIFNSPNVNLSTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAD 399
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
499 QILLEFIEDLNN-PSVSLTPTTFNELVTQRKHNEVMVVDVFSPPWCHPCQVLMPEKRMAR 557
```

```

Qy 400 KLAGSTGMKVGKFRADGGQKFAKSELOLGSFPTILFFPKHSR--PTIKYPSEKRDVDS 457
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 558 TLTGL--INVGSi--DCQYHSFCAQENVQRYPEIRFFPKSNKAYQYHSYNGWNRDAYS 613
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 458 L 458
      |
Db 614 L 614

RESULT 4
US-10-821-234-1287
; Sequence 1287, Application US/10821234
; Publication No. US20050255114a1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1287
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1287

```

RESULT 5
US-10-821-234-1282
; Sequence 1282, Application US/10821234

```

; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1282
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1282

Query Match      4.8%; Score 117; DB 6; Length 440;
Best Local Similarity 27.0%; Pred. No. 0.011;
Matches 30; Conservative 14; Mismatches 53; Indels 14; Gaps 1;

Qy 332 GLSQSHANGDATVVPDIFNSPNVNLISRTGIENLAKLEDKKEPWLVLVLYAPWCPYCOAME 391
Db 152 GRSDSSSKDVIELTDDSPDKNLV-----DSEDVMMVEFYAPWCGCHCKNLE 197

Qy 392 ESYVDLADKLASGTGMKVGKFRADGGEQKEFAKSELQLGSPPTILFFPKHSS 442
Db 198 PEWAAAASEVKEQTGKGKLAADVATVNVQLASRYGIRGPTTIKIFQKGES 248

RESULT 6
US-10-878-556A-114
; Sequence 114, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 114
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: tr_hum/q9bv9h
; DATABASE ENTRY DATE: 2001-06-01
US-10-878-556A-114

Query Match      4.5%; Score 110.5; DB 6; Length 324;
Best Local Similarity 27.2%; Pred. No. 0.027;
Matches 41; Conservative 29; Mismatches 68; Indels 13; Gaps 7;

Qy 320 VKQKEEDVN--NGNLSQSHANGDATVVPDIFNSPNVNLISRTGIENLA----KLEDKKE 373
Db 41 VKYQGRPDFQTLERNWMLQTLNEBFTVPEPEV--EPPSAPEL--KQGLVELSASNLFELHVAQG 98

Qy 374 PMLVLVYAPWCPYCOAMEESYVDLADKLASGTGMKVGKFRADGGEQKEFAKSELQLGSPPT 433
Db 99 DHFIKFPAPWCGCHCKALAPTWEQALGLEHSEVTVKGK--VDCTQHYELCSGNQVRGYPT 156

Qy 434 ILFFPKHSSRPPTIKYPSEKRDVDSLMAFVNA 464
Db 157 LLNWF--RDGKKVDQYKG--KEDLESLEYVES 184

RESULT 7
US-11-000-463-344

```


Db 87 IKFYAPRGCHCKTAPTWEELESKEFFGLAGVKIARVDCTAERNICKYSVR--GYPTLL 144
Qy 436 FFKPHSSRPITIKYPSKRDVDSLMAFV 462
Db 145 LF---RGKKVSEHSGRDLDLSLRFV 168

RESULT 10

US-10-878-556A-1
; Sequence 1, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (307)..(307)
; OTHER INFORMATION: Xaa is an undetermined amino acid residue
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: hsupg:006101-3-0
; DATABASE ENTRY DATE: 2003-02-16
US-10-878-556A-1

Query Match 4.2%; Score 101.5; DB 6; Length 414;
Best Local Similarity 22.6%; Pred. No. 0.24; Mismatches 34; Indels 47; Gaps 11;
Matches 50; Conservative 34;
Qy 250 WNPVANYKGHDIMFLRTMNPVNSLHAKGVSVIGCE-----PCTRPVLPQGHREGR 302
Db 119 WNDLIG----DKYNSMEDAKVYVAKVDTAHSV--CSAQGVGYPTLKLKFKPGQEA---- 168
Qy 303 WWEDAKAKEGLHKNVKKQKEEDVN--GNGLSQSHANGDATVDFINSNPVNLRT 360
Db 169 -----VKYQGRDFQTLNWMMLQTLNEEPVTPPEPV--BPPSAPEL--KQ 209
Qy 361 GIENLA----KLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVKGFADG 416
Db 210 GLVELSASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQALGLEHSETVKIGK--VDC 267
Qy 417 EQKEFAKSELQGSFPTILFFPKHSSRPITIKYPSKRDVDS 457
Db 268 TQHYELCSGNQVRGYPTLLWF--RDGKKVDQYKG--KRDLES 305

RESULT 11

US-11-024-959-352
; Sequence 352, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30

; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 352
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-352

Query Match 4.0%; Score 99; DB 7; Length 875;

Best Local Similarity 22.0%; Pred. No. 1.2;
Matches 100; Conservative 52; Mismatches 147; Indels 156; Gaps 27;
Qy 112 AIAFSGAEDVAL-----IEYAK-----LTG-----RPFVVF-----SLDT 141
Db 63 SVAFDGAEVLVLGASSGVIKLWDVEAKMVRGLTGHRSNCTAMEFHPFGEFFASGSTD 122
Qy 142 GRLNPETYQLPDAVEK-----HYGRIEYMPD-----AVEVQALVRSKG 181
Db 123 -----NLKIWDIRKKGCIHTYKGTGTGISTIRFSPDGRVWVSGNDNVVVKWDLTAGKL 176
Qy 182 L--FSFYEDGHQECRCVRKV--RPLRRAL-----KGLRAW-----ITQQRKDQSPGTR 225
Db 177 LHDQFKHEN-----HIRSIDFHLFELLATGSDARTVKFWDLETTELIGSSRPAAGVR 230
Qy 226 SEIPVVOVDP-----AFEGMDGGIGSLVKMNPVANYKGHDI--WNPLRTMNPVNSLHAKGY 280
Db 231 A-----IAFHPDGRTLFCGLEDSL--KVYSWEPIVCHDGVDMGWSTLADLCIHGDKL----- 280
Qy 281 VSIQCEPCTRPVLPQGHREGRWWEDAKAKEGLHKNVKKQKEEDVNGLSQSHANG 340
Db 281 --LGC-----SYIQSGVGVVADASLIE--PYGTNYKPO--QKDSGDEIEHQSRP 326
Qy 341 DA---TTV-----PDIFN-----SPNVNLSRTGIENLAK-----LDRK 372
Db 327 SAKVGTIRSTIMRCASPDYETKDKNIYVDTASGNPVSSQVRVTFNFAKVTQPLDFND 386
Qy 373 EPMVLVLYAPWCPYCOAMEESYVDLADKLAGSTGMK-----VGKPRADGEQKEFAKSE 425
Db 387 TPNLTL-----RRQGLVTETPDGLSGHVPSKSIQPKVVSRRSDPDGKDSRRRESI 436
Qy 426 LQIGSEPTILFFPKHSSRPITIKYPSKRDVDSLMA 460
Db 437 TFSRTKPGMLLRPAHSRRPS-----STKYDVDRLSA 467

RESULT 12

US-10-467-962B-6
; Sequence 6, Application US/10467962B
; Publication No. US20050246784A1
; GENERAL INFORMATION:
; APPLICANT: Plesch, Gunnar
; APPLICANT: Blau, Astrid
; APPLICANT: Daeschner, Klaus
; APPLICANT: Klein, Mathieu
; TITLE OF INVENTION: Identification of Herbicidally Active Substances
; FILE REFERENCE: 2000 857
; CURRENT APPLICATION NUMBER: US/10/467,962B
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: PCT/EP02/01466
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patent in Vers. 2.0
; SEQ ID NO 6
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-467-962B-6

Query Match 3.7%; Score 91.5; DB 6; Length 183;

Best Local Similarity 31.9%; Pred. No. 0.57;
Matches 30; Conservative 14; Mismatches 45; Indels 5; Gaps 4;

Qy 354 VVNLRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVKFR 413

Db 77 VKLSAQLGELVK-GRKVLPLVDYATWCGPCILMAQLEMLA--VEYESNAIIKVD 133

Qy 414 ADGEQKEFAKSELQGLSFPTILFPKHSRPTIK 447

Db 134 TDDEY-EFAR-DMQVRLPTLFFISPDPSKDAIR 165

RESULT 13

US-10-763-712A-48

; Sequence 48, Application US/10763712A

; Publication No. US20050266541A1

; GENERAL INFORMATION:

; APPLICANT: Solazyme, Inc.

; APPLICANT: Dillon, Harrison F.

; TITLE OF INVENTION: Methods and Compositions for Evolving Microbial Hydrogen

; TITLE OF INVENTION: Production

; FILE REFERENCE: H2042101-CIP

; CURRENT APPLICATION NUMBER: US/10/763,712A

; CURRENT FILING DATE: 2004-01-21

; PRIOR APPLICATION NUMBER: US 10/287,750

; PRIOR FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: US 10/411,910

; PRIOR FILING DATE: 2003-04-12

; PRIOR APPLICATION NUMBER: US 60/500,032

; PRIOR FILING DATE: 2003-09-03

; NUMBER OF SEQ ID NOS: 184

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 48

; LENGTH: 844

; TYPE: PRT

; ORGANISM: Desulfitobacterium hafnienae

; US-10-763-712A-48

Query Match 3.7%; Score 91.5; DB 6; Length 844;

Best Local Similarity 19.7%; Pred. No. 5.4;

Matches 84; Conservative 52; Mismatches 142; Indels 149; Gaps 18;

Qy 1 MALAFTSSISAPTSFTFSSEPKLPQIGSIRISERPIGAVNNLSQRRSLVKPVNAEPPR 60

Db 240 LSAAYSAIQHVDVTVERHPR-----AGGMRYGIPYRLPKETLDRE--- 283

Qy 61 KDSIVPLAA-----TTIVASASETKEDPQIASDLD-----NASPLEIMDRALD 105

Db 284 ----IGLTADLVGKIMTKALGTHIRLEDLHQ-----DFDAVLAIGSWRATPLQTEGDNLE 336

Qy 106 -----KFGNDIAIAFSG--ABEDVALIEYAKLTGRPFVFSLDTGRLN 145

Db 337 GWLGINFLEQVTKGADIKLGEHVVVIGGGNTAIDCARTALKAGSVKLVYRRTREMP 396

Qy 146 PETYQLFDVAKHYGIRIEYMFDPDAVEQALVRSKGLFSFYEDGHQECRCVKVRPLRA 205

Db 397 AESYEVEBEAI--HEGVEMYFL-----TAPHKIVABEGG-----RKLHLC 432

Qy 206 LKGLRAWITGQRKQDSQPSCTSEIPVQVDPAPFEG--MDGGIG-----SLVK 249

Db 433 IK-----MTLGEDRS--GRRPIPIEGSETAFEDATTIIGAIGQSTNTQFLVHDLPVKLNK 486

Qy 250 WNPVANVKGHDIMNLFRTMNVVNSLHAKGVYSIGCEPCTRPVLPGQHERGRWWEDAK 309

Db 487 WGD I-EING-----KTMQISENNIPAGGDCVTGPATVIQAAAGRA-----AE 529

Qy 310 AKEGCLHKNVYKQKEEDVNGLSQSHANGDATVVDI FNSPNVNLSPRTGIENLAKLE 369

Db 530 AMSDFLAKMGYYKEQ-----PMDYSCSRGSLDLPQWE 561

Qy 370 DRKEPWL 376

Db 562 FEKIPRL 568

RESULT 14

US-11-156-084-227

; Sequence 227, Application US/11156084

; Publication No. US20060010515A1

; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC

; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to

; TITLE OF INVENTION: agronomically interesting phenotypes

; FILE REFERENCE: (38-21)

; CURRENT APPLICATION NUMBER: US/11/156,084

; CURRENT FILING DATE: 2005-06-17

; NUMBER OF SEQ ID NOS: 364

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 227

; LENGTH: 310

; TYPE: PRT

; ORGANISM: Clostridium perfringens

; US-11-156-084-227

Query Match 3.7%; Score 90; DB 7; Length 310;

Best Local Similarity 23.2%; Pred. No. 1.7;

Matches 59; Conservative 46; Mismatches 81; Indels 68; Gaps 16;

Qy 22 KLPQIGSIRISERPIGG-----AVNFNLSQRRSL-----VKPVNA--EPKRD SI 64

Db 41 KYMDIGSAKISKEMGGIPHYLIDFVDPSPKESVAFKDLTTEKIKDIQSRGKLPILVGG 100

Qy 65 VPLAATTIVAS--ASETKEEDF-----EQIASDLDNASPLEIM-DRALDKFGNDIAIAFS 116

Db 101 TGLYINSLICNMNPAESDKDESYRELEKIANEHNGLYHLEMLKDIDLESYN---SIHFN 157

Qy 117 GAEDV-ALIEYAKLTGRPFVFSLDTG-----RLNPETYQLFDAVEKHGIRIE 164

Db 158 NRKRVIRALETYKLTGKPFSSFRKAKNSIYETPNYIVYVLMMDRAKLYDRINK---RVD 213

Qy 165 YMFDPDAV--BVOALVRSKGLFS-----FYEDGH---QSCCRVKVRPLRA 205

Db 214 IMFEKGLLEEVKNL-KAMGLTDDMQSMKGICYKEVLYLDGKISLEQCIEMIK-QGSRNY 271

Qy 206 LKGLRAWITGQRKD 219

Db 272 AKRQLTWTF---RKD 282

RESULT 15

US-10-858-730-76

; Sequence 76, Application US/10858730

; Publication No. US20050255568A1

; GENERAL INFORMATION:

; APPLICANT: Blomquist, Richard B.

; APPLICANT: Bailey, Paul

; APPLICANT: Doten, Reed

; APPLICANT: Driggers, Edward M.

; APPLICANT: Madden, Kevin T.

; APPLICANT: O'Leary, Jessica

; APPLICANT: O'Toole, George

; APPLICANT: Trueheart, Joshua

; APPLICANT: Walbridge, Michael J.

; APPLICANT: Yorgey, Peter S.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID

; FILE REFERENCE: 14184-030001

; CURRENT APPLICATION NUMBER: US/10/858,730

; CURRENT FILING DATE: 2004-06-01

; PRIOR APPLICATION NUMBER: US 60/475,000

; PRIOR FILING DATE: 2003-05-30

; PRIOR APPLICATION NUMBER: US 60/551,860

; PRIOR FILING DATE: 2004-03-10

; NUMBER OF SEQ ID NOS: 364

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 76

; LENGTH: 760

; TYPE: PRT

; ORGANISM: Mycobacterium leprae

; US-10-858-730-76

Query Match 3.7%; Score 89.5; DB 6; Length 760;
Best Local Similarity 19.3%; Pred. No. 7;
Matches 110; Conservative 76; Mismatches 212; Indels 173; Gaps 25;

Qy	13	TSTPPSEPKLPQIGSTRISERPIGG--AVNFNLQRRSLVKPVNAEPKDSIVFLAAT	70
Db	7	TQSFTATVVTGSPRIGPRRELKATEGYWAKTSRSELESV-----ASTLRDMWSDIAA-	60
Qy	71	TIVASASETKEEDEFQIATSDNASPLEIMDRALDKFGNDIAIAFS---GAEDVALIEYA	127
Db	61	---AGLDSVPNTFSYYDQMDLTAFMLGALPARVAQVSDDDLDQYFALARGNNDIKPLEMT	117
Qy	128	KLTC-----RPRVFSLDTGRNLNPETYQLFDAVEKHYGIRIEYMPFDDAVEQALV	177
Db	118	KWFDTNVHYLVPEIEPATTFSLNPKILG--LKEALEQRIPSRPVIIGP---VTFL	170
Qy	178	RSKGLFSFYEDGHECCRVKVRPLRRALKGLRA-----WITGQ-----KQSP-----	222
Db	171	LSKGI-----NGGGAPIQRLBELVGIYCTLLSLAENGARWVQFDEPALVTDLSPDAPALA	226
Qy	223	-----GTRSEIPVVQVDPAPFEGMDGGIGSLVKWNEV-----	253
Db	227	EAVYTALGVSXKRPATYVATYFGNPGASLAGLAR-TPIEAIGVDFVCGADTSVAAPPELA	285
Qy	254	-----ANVKGHDIV-----NFLRTMNVPN-----SLH	276
Db	286	GKTLVAGIVDGRNIWRTDLBSALSKLATLGSAAATVAVSTSCSTLHVPSLEPETDLDN	345
Qy	277	AKGVVSGCEPCTRPVLPQOHEREGRWMDAKAKECG-----LHKGNVKQ	322
Db	346	LRSLAFGAEKVAEVVVLARALRDGR---DAVADEIAASNAAVASRRSDPRLHNGQVRA	401
Qy	323	QKEEDV-----NGNLSQSHANGDA-----TTVPDIFNSPNVNLSTGTIENAKLED	370
Db	402	RIDSIVASGTHRGDAQRRTSQDARLHLPLPTTTIGSYPTQSAIRKARAALQD-AEIDE	460
Qy	371	RKEPWLVLVYAPWCPYQAMEESYVDLADKLAGSTGMKV---GKFRADGQKEFAKSELQ	427
Db	461	AE-----YISRMKEVAD-AIKLQSLGLDLVLHGEPENDMVQYFAE---Q	503
Qy	428	LGSF-----PTILFFPKHSSRPTIKYPSEKR	453
Db	504	LGGFFATQNGWQSGSRCVRPPILYGDVSR	534

Search completed: February 18, 2006, 14:52:16
Job time : 18 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n-model

Run on: February 20, 2006, 22:02:50 ; Search time 7245 Seconds
(without alignments)
3656.183 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPSTTPFSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 2842125653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB.spool/US10731525/runat.17022006.164857.82/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=ptt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HCP=abes05h
-USER=US10731525 @CGN 1.1 4939 @runat.17022006.164857.82 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2445	100.0	1629	6 AR533623
2	1951	79.8	1705	15 AF452450
3	1923	78.7	1795	6 AR533622

4	1883.5	77.0	1648	15	AY353089	AY353089 Populus x
5	1832.5	74.9	1386	15	AY568717	AY568717 Lycopersi
6	1832.5	74.9	1835	15	BT013254	BT013254 Lycopersi
7	1829	74.8	1753	15	CRU63784	U63784 Catharanthu
8	1811.5	74.1	1398	6	CQ805882	CQ805882 Sequence
9	1811.5	74.1	1398	6	AX412862	AX412862 Sequence
10	1811.5	74.1	1398	6	AX505803	AX505803 Sequence
11	1811.5	74.1	1398	6	AX651565	AX651565 Sequence
12	1811.5	74.1	1398	15	BT002612	BT002612 Arabidops
13	1811.5	74.1	1710	15	AF424582	AF424582 Arabidops
14	1808	73.9	1608	15	SCJ001208	AJ001208 Brassica
15	1806.5	73.9	1714	15	AY088011	AY088011 Arabidops
16	1801	73.7	1702	15	AY087591	AY087591 Arabidops
17	1799	73.6	103129	15	AP004494	AP004494 Lotus cor
18	1797	73.5	1377	15	AY103313	AY103313 Arabidops
19	1797	73.5	1555	15	AY093319	AY093319 Arabidops
20	1797	73.5	1654	15	AY062665	AY062665 Arabidops
21	1797	73.5	1709	15	AF428445	AF428445 Arabidops
22	1797	73.5	1712	15	AY054175	AY054175 Arabidops
23	1795.5	73.4	1748	15	ATU53864	U53864 Arabidopsis
24	1784	73.0	1763	15	ATU53865	U53865 Arabidopsis
25	1758.5	71.9	1687	15	AY088665	AY088665 Arabidops
26	1758.5	71.9	1712	15	ATU96045	U96045 Arabidopsis
27	1745.5	71.4	1945	15	AF016282	AF016282 Arabidops
28	1745.5	71.4	47319	6	AX059533	AX059533 Sequence
29	1745.5	71.4	103632	15	AF074021	AF074021 Arabidops
30	1745.5	71.4	200001	15	ATCHRIV13	AL161501 Arabidops
31	1738	71.1	1365	6	CQ805458	CQ805458 Sequence
32	1738	71.1	1365	6	AX412892	AX412892 Sequence
33	1738	71.1	1365	6	AX412893	AX412893 Sequence
34	1738	71.1	1365	6	AX507610	AX507610 Sequence
35	1738	71.1	1365	15	AY040005	AY040005 Arabidops
36	1738	71.1	1783	15	AF360192	AF360192 Arabidops
37	1735	71.0	1365	15	AF023167	AF023167 Arabidops
38	1734	70.9	1634	15	ATU43412	U43412 Arabidopsis
39	1728.5	70.7	1738	15	ATU53866	U53866 Arabidopsis
40	1717	70.2	1482	15	ATU56922	U56922 Arabidopsis
41	1715	70.1	1700	15	SCJ001207	AJ001207 Brassica
42	1690.5	69.1	2945	15	AF016284	AF016284 Arabidops
43	1687	69.0	103270	15	ATFLN20	AL022140 Arabidops
44	1687	69.0	196247	15	ATCHRIV56	AL161556 Arabidops
45	1668.5	68.2	1809	15	AY739296	AY739296 Zea mays

ALIGNMENTS

RESULT 1	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
DEFINITION	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
ACCESSION	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
VERSION	AR533623.1	GI:539233592	1629 bp	DNA	linear	PAT 08-OCT-2004
KEYWORDS	AR533623.1	GI:539233592	1629 bp	DNA	linear	PAT 08-OCT-2004
SOURCE	Unknown.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
ORGANISM	Unknown.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
REFERENCE	1 (bases 1 to 1629)	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
AUTHORS	Falco, S.C.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
TITLE	Genes encoding plant adenosine 5'-phosphosulfate reductase	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
JOURNAL	Patent: US 6730827-A 7 04-MAY-2004;	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
FEATURES	E. I. du Pont de Nemours and Company; Wilmington, DE	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
source	Location/Qualifiers	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
ORIGIN	1..1629	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
	/organism="unknown"	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
	/mol_type="genomic DNA"	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
Alignment Scores:	6.49e-175	Length:	1629			
Pred. No.:	2445.00	Matches:	466			
Score:	100.0%	Conservative:	0			
Percent Similarity:	100.0%	Mismatches:	0			
Best Local Similarity:	100.0%	Indels:	0			
Query Match:	100.0%					

DB:	6	Gaps:	0
US-10-731-525-8 (1-466) x AR533623 (1-1629)			
Qy	1	MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu	20
Db	32	ATGGCCCTCGCTTCACTTCTCAATTTCCGCACCACTTCCACCTCCCATCATCGAA	91
Qy	21	ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyAlaVal	40
Db	92	CCCAAACTTCGCAAAATGGGTCAATTAGGATTTCCGAGAGGCCCAITGGAGGCGCGTT	151
Qy	41	AsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluProProArg	60
Db	152	AATTTCAATTTATCTCAAGACCGAGCTTGGTAAGCCCGTTAAACGCCGAACCTCCACGC	211
Qy	61	LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys	80
Db	212	AAGGATTCATTTGCTCTCGCAGCAACAACCATCGTTGCTTCTGCTCTGAGACGAA	271
Qy	81	GluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet	100
Db	272	GAGGAAGATTTGAACAGATAGCCAGTGATCTGCACAAATGCTTCACTCTTGAATCATG	331
Qy	101	AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp	120
Db	332	GATAGACCCCTCGCAAAATTCGGACAGCATAGCTATTTGCTTTCAGTGGTCTGAAGAT	391
Qy	121	ValAlaLeuIleGluTyrrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp	140
Db	392	GTTCCTTTGATTTGATGATGCAAAATTCAGCGGTGACCCCTTTAGGTTTTCAGTTTGGAC	451
Qy	141	ThrGlyArgLeuAsnProGluThrTyrrGlnLeuPheAspAlaValGluLysHisTyrrGly	160
Db	452	ACTGGAGAGTGAACCCAGAAACTTATCAACTTTTGTATGCGGTGAGAAACATATGGA	511
Qy	161	IleArgIleGluTyrrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys	180
Db	512	ATTGCGATTGATGATACATGTTCCCTGATGCTGTTGAGGTTTCAGGATTTGAGGAGTAAG	571
Qy	181	GlyLeuPheSerPheTyrrGluAspGlyHisGlnGluCysArgValArgLysValArg	200
Db	572	GGGTATTTCTTTCTACGAGATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGTGAGG	631
Qy	201	ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln	220
Db	632	CCTTTAAGGAGGCGCCTTAAGGGTCTCAGAGCATGGATAACTGGTTCAGAGGAAAGACCG	691
Qy	221	SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet	240
Db	692	TCACCTGGTACTAGGTCTGAAATACCGGTTGTTGAGGTTGATCCGGCTTTTGGAGGAATG	751
Qy	241	AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp	260
Db	752	GATGGTGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAATGTGAAGGGCCATGAC	811
Qy	261	IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrr	280
Db	812	ATATGGAACCTTCTTAGGACCATGAATGCGCTGTGAATTCCTTGCATGCAAAAGGATAT	871
Qy	281	ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu	300
Db	872	GTTTCCATTTGGTGTGAGCCCTGCACTAGGCGCTGTTTACCTGGGCAACATGAAGGGNA	931
Qy	301	GlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal	320
Db	932	GGGAGGTGGTGGGAGGATGCCAAACTTAAGGAATCTGCTTCAACAAAGGAAATGTA	991
Qy	321	LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly	340
Db	992	AAGCAGCAGAAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGCAAAATGGT	1051
Qy	341	AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr	360

Db	1052	GATGCTACCACTGTGCTGACATTTTCAACAGCCCAAGTGTAGTTAATCTTGAGCAGGACT	1111
Qy	361	GlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProThrLeuValValLeuTyrr	380
Db	1112	GGAAATTGAGAATTTGGCAAAATTTGGAGGACCGAAGAACCATGGCTTGTGTCTTTAT	1171
Qy	381	AlaProTrpCysProTyrrCysGlnAlaMetGluGluSerTyrrValAspLeuAlaAspLys	400
Db	1172	GCACCAATGGTGCCCTACTGTCAGGCTATGAGGAATCTTATGTTGACTTAGCAGACAAG	1231
Qy	401	LeuAlaGlySerThrGlyMetLysValGlyPheArgAlaAspGlyGluGlnLysGlu	420
Db	1232	TTAGCAGGGTCAACAGGATGAAGTTGGAAATTTAGAGCAGATGCGAACAAGAA	1291
Qy	421	PheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHis	440
Db	1292	TTTGCAAAGAGTGAATGCAATTTGGGAAGCTTCCCTACGATATTTATTTTCCCAAGCAT	1351
Qy	441	SerSerArgProThrIleLysTyrrProSerGluLysArgAspValAspSerLeuMetAla	460
Db	1352	TCGTCTCGGCCAACATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGATGCA	1411
Qy	461	PheValAsnAlaLeuArg	466
Db	1412	TTTGTAAATGCCTTAAGA	1429

RESULT 2
AF452450 1705 bp mRNA linear PLN 21-JAN-2002
LOCUS
DEFINITION Glycine max adenosine 5'-phosphosulfate reductase mRNA, complete cds

ACCESSION AF452450
VERSION AF452450.1 GI:18252503

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 1705)

AUTHORS Phartiyal, P. and Krishnan, H.B.

TITLE Nucleotide sequence of a soybean cDNA encoding adenosine

5'-phosphosulfate reductase

UNPUBLISHED

REFERENCE 2 (bases 1 to 1705)

AUTHORS Phartiyal, P. and Krishnan, H.B.

TITLE Direct Submission

JOURNAL Submitted (28-NOV-2001) USDA-ARS, University of Missouri, 108W

Curtis Hall, Columbia, MO 65211, USA

FEATURES Location/Qualifiers

source 1..1705

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

16..1428

/codon_start=1

/product="adenosine 5'-phosphosulfate reductase"

/protein_id="AAL66290.1"

/db_xref="GI:18252504"

/translation="MALAVSTNTSSSSAAAAAASSFFSRIGSSSDAKAQAIGSRFRPPER

SLVSVVVNTQRRSLVRPLNAEPQRNDSIVPLAATIVAPVEKEEDFEQIADKLEN

ASPLEIMDRALKEFGNDIATAFSGAEDVALIEYAHILGRPYRPSLDLGRNLNPETYKF

FDAYEKGTHIEYMFDAVEQALVRKGLFSFYEDHQECCKRVRKVRPLRRADKGL

KAWITGQRKDQPTSRSEIPVQVDPFEGLDGGISLVKNPNVNVNGLDIWSFLRT

ADLVNGLHSQGVYSIGCEPTPLVPCOHEREGWWMDAKACEGLHKNIKHEDA

MQLVNGASQANGSATVADIFNSODVVSLSRSGIENLAKENRKEPLVLVYAPWCRF

QAMNESYVDLAEKLAGSVKAKFRADGQKEVAKTELQGLSPPTILLFPKHSSQPI

KYPSEKRDVDSLTAFAVNALR"

ORIGIN

Alignment Scores:

```

Pred. No.: 1.21e-137 Length: 1705
Score: 1951.00 Matches: 382
Percent Similarity: 88.2% Conservativity: 38
Best Local Similarity: 80.3% Mismatches: 40
Query Match: 79.8% Indels: 16
DB: 15 Gaps: 7

US-10-731-525-8 (1-466) x AF452450 (1-1705)

QY 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThr 15
Db 16 ATGGCTCTCGCGTTTCCACTACTTCTCTCTCAGCAGCTGCTGAGCTTCTAGCTCC 75
QY 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 76 TTCTTCTCTGCTCGGATCTTCTCGGAGCGAAAGCTCTCAATCGGTTCGTTTCGG 135
QY 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
Db 136 TTCCGAGAGGTCGCTGTTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 195
QY 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 196 GTGAGCCACTCAACGCCGACCGCAACGACGATTCATTTCTTCTTCTTGAGCA--- 252
QY 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 253 ACTATCGTCTCTCGAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312
QY 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 313 CTTGAAATGCAATCTCTCTCTGAAATTTATGATGAGGCGCTCGAGAAATTTGGGAACG 372
QY 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 373 ATCGTATTGCAATTTAGTGTGCTGAGATGTTGCTTTGATGAGTATGCACATTTGACG 432
QY 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 433 GGTGACCCCTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAA 492
QY 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 493 TTTTGTGATCGGTTGAGAGCAATTTAGGAATTTACATTTGAGTACATGTTCCCTGATGC 552
QY 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 553 GTTGAGTTTCAGGCATTTAGTAAAGCACTAAAGGCTCTTCTCAITTTACGAGGATGGCAT 612
QY 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 613 CAAGAGTGTCTCGAGTAAGGAGGTGAGGCGCTTGAGGAGAGCCCTTAAGGCTCTCAA 672
QY 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 673 GCATGATCACTGGACAGAGAAAGATCAGTCTCTGCTGTAAGTCTGAAATCCCTGTT 732
QY 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 733 GTCAGGTGTATCCTGTTTGTAGGAGCTGGATGGTGGAAATGGCAGCCTGGTGAAGTGG 792
QY 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 793 AACCCGGTTGCAATGTAATGCTTAGACATATGAGCTCTCTTAGGACCATGGATGT 852
QY 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 853 CCTGTAATTCATTGCAATTTCCCAAGGATATGTTTCGATTTGGCTGTGAGCGCTGCAAGG 912
QY 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 913 CCAGTTTTACCGGCTCAACATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 972

```

```

QY 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db 973 AAGGAGTGTGCTTCTCACAAAGGTATATATAACATGAAGATGCTGCCAGCTTAATGA 1032
QY 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1033 AATGGGGCCCTCCCAA-----GCAAAATGGATCTGCC---ACTGTCTGCTGACATTTTCAAC 1083
QY 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1084 TCCAGGATGTGTGTCAGCTTGAGCAGGTCTGGGATTTGGAATTTGGCAAAATTTGGAGAAC 1143
QY 371 ArgLysGluProTrpLeuValLeuValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1144 CGAAAGAACCCCTGGCTTGTGCTCTATGACCACTGGTGGCGCTTCTGTGAGGCTATG 1203
QY 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1204 GAGGAATCATATGTGATTTGGCAGAAAGTTAGCTGGGTCA---GGAGTGAAGGTTGCA 1260
QY 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1261 AAATTAGAGCGGATGGAGACAGAGGAATATGCAAAAGCTGAACTGCACTGCGAGTGG 1320
QY 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1321 TTCCCCACAATACTACTCTTCTCCCAACACTCTTCTCAACCA---ATTAGTACCTTTCG 1377
QY 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1378 GAAAGAGAGATGTTGATTCATTTGAGCGCATTCGTGAATGCTTTACGG 1425

RESULT 3
LOCUS AR533622 1795 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6730827.
ACCESSION AR533622
VERSION AR533622.1 GI:53923591
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1795)
AUTHORS Falco,S.C.
TITLE Genes encoding plant adenosine 5'-phosphosulfate reductase
JOURNAL Patent: US 6730827-A 5 04-MAY-2004;
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
Source 1..1795
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 1.67e-135 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservativity: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x AR533622 (1-1795)
QY 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTTGCGGTTTCCACTACTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 115
QY 16 Phe-----ProSerSerGluProLysLeuProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 175
QY 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50

```

```
Db 176 TTTCGGAGAGGCTCAAGTTTCGTCTGGTGTTGTTAAATTAACCAAGACGCTCCTCG 235
Qy 51 ValLysProValAlaSerGluProProArgLysAspSerIleValProLeuAlaAAlaThr 70
Db 236 GTGAGGCACTCAATGCGCAACCGCAACGGAATGATTCTGTGTCTCTTGCAACA-- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCTGTTCTCTGAGGTGAGAGGAGAAAGAAATTTGAGCAATTTAGGCAAGAC 352
Qy 91 LeuAspAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAATTCATCTCTCTTGAGATTATGATAAGGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThr 130
Db 413 ATCGCTATTGCTTTAGTGTGCTGAAGATGTTGCTTTGATTGAGTATGCACATTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 473 GGTGACCCCTACAGAGTGTTAGTCCTTGACACTGGGAGACTGAACCCAGAAACCTCAAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluThrMetPheProAspAla 170
Db 533 TTTTTCACGCTGTTGAGAGCATTATGGAATTCATATTGAGTACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheThrGluAspGlyHis 190
Db 593 GTTGAGGTTGAGGCATTAGTAAGAACTAAGGGGCTCTCTCAATTTACGAGGATGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGCTGTAGATAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGATTACTGCACAGAGAAAGACAGCTCTCTGCTAGTACTAGTCTGAAATCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 773 GTCCAGGTTGATCTGTTTGTAGGGACTGGATGTTGGAATTTGGCAGCCTGGTGAAGTG 832
Qy 251 AsnProValAlaSerValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGTTGCAATGTTAATGGTCTAGACATATGGAATCTCTCTAGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CTTGTAATTCATTGCTATCCCAAGGATATGTTTCGATTGGCTGTGAGCCATGCACAAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 953 CCGGTTTTTACCGGCACAACATGAAGAAGAAAGGTGGTGGTGGGAGGATGCCAAGCC 1012
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1013 AAGGAGTGTGCTCTTCAAAAGGTAATTTGAAAAGGAAGATGTCGCCAGCTTAATGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAATGGCTCTGCC---ACTGTTGTGTGACATTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGTGTCAGCTTGCAGAGGTCCGGGATTGAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTrpLeuValIleuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAAACACTGGCTTGTGTGCTCTATGCACCATGTCGCCGCTTCTGTGAGCCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1244 GAGGAGTCGTATGTTGATCTGCGCAGAGAAGTTAGCAAGGTCA---GGAGTGAAGTTGCA 1300
```

```
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1301 AAATTCAGAGCCCATGAGAGCAGAGGAATATGCAAAAGAGTGAACCTGCGATTGGGAAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1361 TTCCCAACAATACTTCTCTTCCCAAGCACTCTTCTCAACCA--ATTAAGTACCCCTTCA 1417
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1418 GAAAGAGAGATGTTGATTCATTCAGCGCATTCGTGAATCGCTTACGG 1465
```

RESULT 4

LOCUS AY353089 1648 bp mRNA linear PLN 12-JAN-2005
DEFINITION Populus x canescens adenosine 5' phosphosulfate reductase mRNA,
complete cds.

VERSION AY353089

KEYWORDS AY353089.1 GI:34099826

SOURCE Populus alba x Populus tremula

ORGANISM Populus alba x Populus tremula

Populus alba x Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustosids I; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 1648)
Kopriva,S., Hartmann,T., Massaro,G., Hoenicke,P. and Rennenberg,H.
AUTHORS Regulation of sulfate assimilation by nitrogen and sulfur nutrition
TITLE in poplar trees

JOURNAL Trees (Berl. West) 18, 320-326 (2004)

REFERENCE 2 (bases 1 to 1648)

Hartmann,T., Hoenicke,P., Wirtz,M., Hell,R., Rennenberg,H. and

Kopriva,S.

Regulation of sulphate assimilation by glutathione in poplars

(Populus tremula x P. alba) of wild type and overexpressing

gamma-glutamylcysteine synthetase in the cytosol

J. Exp. Bot. 55 (398), 837-845 (2004)

REFERENCE 15020643

3 (bases 1 to 1648)

Kopriva,S.

Direct Submission

Submitted (28-JUL-2003) Institute of Tree Physiology,

Georges-Koehler-Allee 053, Freiburg 79085, Germany

Location/Qualifiers

FEATURES 1. 1648

source /organism="Populus alba x Populus tremula"

/mol_type="mRNA"

/db_xref="taxon:80863"

/clone="INRA 717 1B4"

74..1471

/codon_start=1

/product="adenosine 5' phosphosulfate reductase"

/protein_id="AA057202.1"

/db_xref="GI:34099827"

/translation="MALSVSSSSISASGFSRSSQELKAPQGSFKLLDRPTTFVNVS

SQRCAVKPLNAEPKNGSVPLAATIAPEIAEKVVEDYDEYKLAELVNGSPLEIM

DKALEKPGDDIAFAEGADVVALEIYAKLTGRPRVFSLDTRLNLPETYHPFDQVEKH

YGRITYEPDPAVEQALVRNKGILFSYEDGHQECRRVKVRPLRRALKGLRAWITGG

LKSGQYTSICEPCTRPVLPGQHERGRWWEDATAKECGLHKGNLKQGDAAQLNGNG

NGAHANGAATVADIFENSENLNLSRPGIENLLKLNKBPWLVLVAPWCQFCQME

ASVDELKLAGSGVKVKFRADGQKGFQKQELQLGSFFTLILFFPKHSSRPKYPSE

KRDVDSLMTFVNALR"

ORIGIN

Alignment Scores:
Pred. No.: 1,43e-132 Length: 1648
Score: 1883.50 Matches: 366
Percent Similarity: 85.3% Conservative: 35
Best Local Similarity: 77.9% Mismatches: 60
Query Match: 77.0% Indels: 9
DB: 15 Gaps: 7

```
US-10-731-525-8 (1-466) x AY353089 (1-1648)
QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPhe-----ProSer 18
Db 74 ATGGCATTTGAGTGTAGTCTTCTCTTCGATTCTCTGCTCTGGGTTTTCACGTTCTAGT 133
QY 19 SerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGly 38
Db 134 CAAGAATCTTAAAGCACCCAGTTTGGTCTCTTTTAAAGCTGTAGATCGGCCGACA--ACA 190
QY 39 AlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluPro 58
Db 191 TTTGTTAATGTTCT--TCACAAAGACGCTGTGCTCTAAAGCCCTTTAAATGCTGAGCCT 247
QY 59 ProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGlu 78
Db 248 AAAAGGAATGGTCTGTGTTCTCTCGCTCGCACTATCGTGTCCAGAGATGCGAGAG 307
QY 79 -----ThrLysGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerPro 96
Db 308 AAAGTAGAGGTGGAGGAAGATTATGAGAAGTTAGCTAAGGAGCTTGTAAATGGTCTCTCT 367
QY 97 LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer 116
Db 368 TTGGAAATATGGAATAGGCACCTTGAGAAATTTGGGGATGACATTTGCTATTGCTTTCCAGT 427
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
Db 428 GTGTCGCGAGATGTTGTTGATTGATGATGCAAGTTGACTGGCGCACCGCTTCAGGGTG 487
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu 156
Db 488 TTCTCACTGGATACGGGGCGGTGGAACCCGAGACATACCATTTCTTTTGACCAAGTTGAG 547
QY 157 LysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeu 176
Db 548 AAGCACTATGCGCATTCGAATGAGTACATGTTTCCAGATGCTGTGTAAGTTCGAAGCATTA 607
QY 177 ValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysAspVal 196
Db 608 GTAAGAACCAAGGTTTATTCTCTTTTACAGATGGGATCAGAGTGTCTCGCGGTGTC 667
QY 197 ArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGln 216
Db 668 AGGAAGGTGAGACCTTTGAGCGGGCTCTGAAGGGCTTCGGGCTCGATCACTGGCCAA 727
QY 217 ArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAla 236
Db 728 AGGAAGGATCAATCTCCAGGAGCAGAGTCTGAAATTCAGTTGTTCAGGTGGACCCCGTT 787
QY 237 PheGluGlyMetAspGlyIleGlySerLeuValLysTrpAsnProValAlaAsnVal 256
Db 788 TTTGAGGATTTGATGTGGGGCTGGAAGCCTTATCAAGTGGATCCATGCGCAATGTT 847
QY 257 LysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHis 276
Db 848 GAGGACCAAGATGTGAAAGTTCTCTCGAACTATGATGTGCTGTGAACTCATTGCAT 907
QY 277 AlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGln 296
Db 908 TCAAAGGATACATCTCTATTGTTGCGAGCCTTTCGACAAAGGCCAGTTCTTCTCGGCCAA 967
QY 297 HisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis 316
Db 968 CATGAGAGAGAGACCGTGTGTGGGAGATGCCACAGCTAAGGAATGTGTTCTTCAT 1027
QY 317 LysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSer 336
Db 1028 AAAGGAATCTGAACAGGGTGATCGACCCCACTTATGCGCAATGGAACGGGGCTGCC 1087
QY 337 HisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsnValValAsn 356
|||||
```

```
Db 1088 CATGCAAAATGGG---GCTGCCACTGTTGCTGATATTATTTCAACTCCGAGAAATTTGGTCAAT 1144
QY 357 LeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeu 376
Db 1145 TTGAGCAGGCCCGGAATCGAGAACTTGTAAATATTAGAGAACCGAAGGAACCATGGCTT 1204
QY 377 ValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAsp 396
Db 1205 GTGGTGTCTCTACGCCCATGGTGCCAAATTTTGCAGAGTATGGAAGCATCTATGTTGAA 1264
QY 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416
Db 1265 TTGGCCGATAAGTTAGCAGGAGT---GGAGTAAAGGTGGGAAATTCAGGGCAGATGGC 1321
QY 417 GluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhe 436
Db 1322 GATCAGAGGGGATCTCAAAACAAGAAATTACAGCTAGGAAGTTTTCACACATCTTTC 1381
QY 437 PheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArgAspValAsp 456
Db 1382 TTCCTTAAACACATCTCGCCCA---ATCAAGTACCCCTCTGAGAAAGGAGCGTGGAT 1438
QY 457 SerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1439 TCATTGATGACTTTTGTGAATGCTCTCCGG 1468

RESULT 5
AY568717 1386 bp mRNA linear PLN 24-JUN-2005
LOCUS Lycopersicon esculentum adenylyl-sulfate reductase mRNA, complete cds.
ACCESSION AY568717
VERSION AY568717.1 GI:51457939
KEYWORDS Lycopersicon esculentum (Solanum lycopersicum)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; Jamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1386)
AUTHORS Coker, J. S., Vian, A. and Davies, E.
TITLE Identification, accumulation, and functional prediction of novel tomato transcripts systemically upregulated after fire damage
JOURNAL Physiol. Plantarum 124 (3), 311-322 (2005)
REFERENCE 2 (bases 1 to 1386)
AUTHORS Coker, J. S., Vian, A. and Davies, E.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2004) Botany, North Carolina State University, Gardner Hall, Campus Box 7612, Raleigh, NC 27695, USA
FEATURES
source 1..1386
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
CDS 1..1386
/codon_start=1
/product="adenylyl-sulfate reductase"
/protein_id="AAU03359.1"
/db_xref="GI:51457940"
/translocation="MALTFTSSAIHGSLSSSSSSYEQPKVSQLGTQFQPLDRQLLSS TVLNSRRSAVKPLYPKRNDSIVPSAATIVAPEGVESVEASDFEKALQNASPL EYHDKALEFPGDIALAFSGAEDVALIEYAHLTGRPYRVSFSLDTGRINPETYQLFDTV EYHGRIRIEMFPDPSIOVALVTKGLFDFEDGHOBCCRCRVRPLRRALGLRAMI TGORKDQSPCTRSEIPIVOVDPSEFGLDGGAGSLKWNPVANVDGKDIWNFLAMNVP VNSLHQGVYSIGCEPCTRPVLPQCHERERHWEADAKACEGLHKGNIKDETVNGAA QNGTATVADIFDTKIDIVLSKPVENVLKEIDRRPMLVLYAPWCQPCQAMEGSTV ELAEKLAGSGVKFKPRADGQKAFQAEELQSGSPFTILFFPKHSSKAIKYPSEKRDV DSLIAFVNALR"
transit_peptide 1..213
/notes="chloroplast transit peptide"
/evidence=not_experimental
misc_feature 346..963
/notes="Region: phosphoadenosine phosphosulfate reductase"
```

```
misc_feature      1048..1377
/ note="Region: thioredoxin domain 2"

ORIGIN
Alignment Scores:
Pred. No.:      8 19e-129      Length:      1386
Score:          1832.50      Matches:      355
Percent Similarity: 85.4%      Conservative: 49
Best Local Similarity: 75.1%      Mismatches: 50
Query Match:      74.9%      Indels:      19
DB:              15          Gaps:          8

US-10-731-525-8 (1-466) x AY568717 (1-1386)
Qy 1 MetAlaLeuAlaPheThrSer-----SerIleSerAlaProThrSer 14
Db 1 ATGSCCTTTGACTTTTCACTTCTTCATCTGCAATTCATGGCTCTTTGTCTTCTTCATCTTCT 60
Qy 15 ThrPheProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArg 34
Db 61 TCTTAT-----GAACAACCCAAAGTATCCCAATTTGGGTACCTTTCAGCCCAATTGGATAGG 114
Qy 35 ProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProVal 54
Db 115 CCTCAACTATTGTCGTCACACTGTTTGNATTTCTCGGAGCGGTCGCGAGTGAAGCCATTG 174
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 175 TATGCTGAACCTAAGAGGAATGATCAATAGTTCGTCAGCAGCTACCATGTCGGTCCT 234
Qy 75 SerAlaSerGluThrLysGlu--GluAspPheGluGlnIleAlaSerAspLeuAspAsn 93
Db 235 GAGTAGGAGAGAGTGTGAGCGCAGAGGACTTTGAGAAATTTGGCTAAGGAGCTTCAAAAT 294
Qy 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113
Db 295 GCTTCCCTCTTTGAGGTATTGCAACAAGCAGCTTGAGAAATTTGGAGATGACATTTGCTATT 354
Qy 114 AlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgPro 133
Db 355 GCTTTCAGTGTGCTGAAGATGTTGCTTTGATAGATACGCACATTTAACTGGACGACCA 414
Qy 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAsp 153
Db 415 TACAGAGTATTTCAGCTTGTACTGAGGAGGTGAACCCGGAGACATACCAATTATTTTGAC 474
Qy 154 AlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluVal 173
Db 475 ACAGTGGAGAGCACTATGGCATTCGCATTGAATACATGTTCCCTGATTCAAGTTGAAGTT 534
Qy 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 193
Db 535 CAGCGTTTGGTTAGGACCAAGGGGCTTTCTTCTTCTATGAGGATGGCCACCAAGAGTGT 594
Qy 194 CysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTyrIle 213
Db 595 TGCCGTGTAAAGAAAGGTTAGGCCCTTTGAGGAGAGCTCTAAAGGGGCTTACGGCGCTGGATC 654
Qy 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnVal 233
Db 655 ACAGCCAGCGTAAGATCAGTCCCTCGGAATCGATCAGAAATCCCATTTGTTCAAGTG 714
Qy 234 AspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTyrAsnProVal 253
Db 715 GACCCCTTTTGGGGTTGGATGGCGGTGCTGGTAGCTTGGTGAAGTGAACCCCTGTG 774
Qy 254 AlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProValAsn 273
Db 775 GCTAATGTGGACGGAAAGATATTGGAACTTCCTGCGTGCATGAATGTGCTGTGTAAC 834
Qy 274 SerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeu 293
```

```
Db 835 TCATTGCAATTCACAAGGATATGATATCCATTGGATGCGAACCTTGCACACAGCCCAAGTTCTA 894
Qy 294 ProGlyGlnHisGluArgGluGlyArgTyrTrpTyrGluAspAlaLysAlaLysGluCys 313
Db 895 CCAGGGCAACACAGAGAGAGGAGATGTTGGGAGATGCCAAGGCCCAAGAGTGT 954
Qy 314 GlyLeuHisLysGlyAsnValLysGlnGlnLysGlnGluAspValAsnGlyAsnGlyLeu 333
Db 955 GGCTTGCACAAGGCCAACATC-----AAGGATGAACCTGTAAATGGC----- 996
Qy 334 SerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsnSerProAsn 353
Db 997 ---GCTGCCCAAAACAAATGCTACTGCT---ACCGTTGCTGATATTTTGTATACCAAGGAC 1050
Qy 354 ValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGlu 373
Db 1051 ATTGTTTACCTTGGATAGCTGAGTGTGAGAACCTAGTAAATTTGGAAGACCCGAAGAGAG 1110
Qy 374 ProTyrLeuValValLeuTyrAlaProTyrCysGlnAlaMetGluGluSer 393
Db 1111 CCTTGGCTCGTGTCTTTATGACCTTGTGTCCAATTTGCCAGGCAATGGAAGGATCC 1170
Qy 394 TyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArg 413
Db 1171 TATGTTCAATTTGGCTGAGAAAGTTGGCTGTTCT---GGTGTGAAAGTAGGAAATTCAGG 1227
Qy 414 AlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThr 433
Db 1228 GCAGATGGTACCAGAAAGCATTTGCACAGAAAGATTTGCAGCTTGGCAGCTTCCCTTACA 1287
Qy 434 IleLeuPhePheProLysHisSerArgProThrIleLysTyrProSerGluLysArg 453
Db 1288 ATACTCTTCTCCCAAGACACTCTTCAAG---GCCATTAAAGTACCTTCCAGAGAAGG 1344
Qy 454 AspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1345 GACGTAGACTCCTTGTCTGGCTTTTGTGAATGCTCTCAGA 1383

RESULT 6
BT013254      1835 bp      mRNA      linear      PLN 11-MAY-2004
LOCUS      Lycopersicon esculentum clone 134763R, mRNA sequence.
DEFINITION
ACCESSION      BT013254
VERSION      BT013254.1 GI:47104669
KEYWORDS      FLI_CDNA.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 (bases 1 to 1835)
AUTHORS      Kirkness,E.F., Wang,W. and Vazeille,A.
TITLE      Direct Submission
JOURNAL      Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1835
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="134763R"
/tissue_type="callus tissue"
/note="TCADM70"

ORIGIN
Alignment Scores:
Pred. No.:      1.14e-128      Length:      1835
Score:          1832.50      Matches:      355
Percent Similarity: 85.4%      Conservative: 49
Best Local Similarity: 75.1%      Mismatches: 50
Query Match:      74.9%      Indels:      19
DB:              15          Gaps:          8
```


US-10-731-525-8 (1-466) x BT013254 (1-1835)

QY 1 MetAlaLeuAlaPheThrSer-----SerIleSerAlaProThrSer 14
 Db 93 ATGGCTTTGACTTTTCACTTCTTCATCTGCAATTCATGGCTTTGTCTTCTTCATCTCT 152
 QY 15 ThrPheProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArg 34
 Db 153 TCTTAT-----GAACAACCAAGATATCCCAATTTGGGTACCTTTTCAGCCATTGGATAGG 206
 QY 35 ProIleGlyAlaValAlaValPheAsnLeuSerGlnArgSerLeuValLysProVal 54
 Db 207 CCTCAACTATTGCTCACTGTTTGAATTTCTCGAGGGCTTCGGCAGTGAAGCCATTG 266
 QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
 Db 267 TATGCTGAACCTAAGAGGAATGATTCAATAGTTCCGTGACGAGCTACCATCGTGGCTCCT 326
 QY 75 SerAlaSerGluThrLysGlu---GluAspPheGluGlnIleAlaSerAspLeuAspAsn 93
 Db 327 GAGGTAGGAGAGAGTGTGGAGCAGAGAGACTTTGAGAAATTTGGCTAAGGAGCTTCAAAAT 386
 QY 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113
 Db 387 GCTTCCCTCTTGAGGTTATGGACAAGACACTTGGAAATTTGGAGATGACATTGCTATT 446
 QY 114 AlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgPro 133
 Db 447 GCTTTCAGTGGTGTGAAGATGTGCTTTGATAGATGACGACATTTAACTTGACGACCA 506
 QY 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAsp 153
 Db 507 TACAGAGTATTACAGCTTGATCTAGCGAGGTGAAACCGGAGACATACCAATTATTGAC 566
 QY 154 AlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluVal 173
 Db 567 ACAGTGAGAGAGCACTATGCACTTCGCAATTCGAATGATGCTTCCTGATTCAGTTGAAGTT 626
 QY 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 193
 Db 627 CAGCGTGTGTAGGACCAAGGGCTTTCTTCTTCATGAGATGGCCACCAAGAGTGT 686
 QY 194 CysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIle 213
 Db 687 TGCGGTGTAAGGAAGGTAGGCTTTGAGGAGAGCTCTAAAGGGCTTACCGCGCTGGATC 746
 QY 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnVal 233
 Db 747 ACAGGCGAGCGTAAGATCATGTCCTCGGAACCTCGATCAGAAATCCCATTTGTCAGGTG 806
 QY 234 AspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProVal 253
 Db 807 GACCTTCTTTGAGGGGTTGGATGGCGGTGCTGTAGCTTGTGTAAGTGGAAACCTGTG 866
 QY 254 AlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsn 273
 Db 867 GCTAATGTGACGCGAAAGATATTTGGAACTTCTCGGTGCGCATGATGTCCTGTGAC 926
 QY 274 SerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeu 293
 Db 927 TCATTGCAATTCACAAGGATATGATCCATTGGATGCGAACTTGCACAAAGCCAGTTCTA 986
 QY 294 ProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLysGluCys 313
 Db 987 CCAGGGCCACACGAGAGAGGGAAGATGTTGGTGGGAAGATGCCAAGGCCAAGGAGTGT 1046
 QY 314 GlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeu 333
 Db 1047 GGCTTGCAAGGCGCAACATC-----AAGGATGAACACTGTAATGTC----- 1088
 QY 334 SerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsn 353
 Db 1089 ---GCTGCCCAACAAATGGTACTGCT---ACCGTTGCTGATATTTTGTATCACCAGGAC 1142

QY 354 ValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGlu 373
 Db 1143 ATTGTTACCTTGAAGTAAAGCTGGAGTTGAGAACTAGTAAATTTGGNAGACCGNAGAGAG 1202
 QY 374 ProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSer 393
 Db 1203 CCTTGGCTGCTGTTCTTTATGCACCTTGGTGCCTTTTGGCCAGGCAATGGAAGGATCC 1262
 QY 394 TyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArg 413
 Db 1263 TATGTTGAATTTGGCTGAGAAAGTTGGCTGTTCT---GGTGTGAAGTAGGGAATTCAGG 1319
 QY 414 AlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThr 433
 Db 1320 GCAGATGGTGACCAAGAAGCATTTGCACAGAGAAATGTCAGCTTGGCAGCTTCCCTACA 1379
 QY 434 IleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArg 453
 Db 1380 ATACTCTTCTTCCCAAGCACTCTTCAAAG---GCCATTAAAGTACCTTCAGAGAGAGG 1436
 QY 454 AspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
 Db 1437 GACGTAGACTCTTGTGCTGCTTTTGTGAATGCTCTCAGA 1475

RESULT 7
 CRU63784
 LOCUS
 DEFINITION
 CRU63784
 DEFINITION
 complete cds; nuclear gene for chloroplast product.
 U63784
 VERSION
 U63784.2 GI:12831473
 KEYWORDS
 SOURCE
 ORGANISM
 Catharanthus roseus (Madagascar periwinkle)
 Catharanthus roseus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinaceae; Catharanthus.
 REFERENCE
 1 (bases 1 to 1753)
 Prior, A., Uhrig, J.F., Heins, L., Wiesmann, A., Lillig, C.H., Stoltze, C., Soll, J. and Schwenn, J.D.
 Structural and kinetic properties of adenylyl sulfate reductase from Catharanthus roseus cell cultures
 Biochim. Biophys. Acta 1430 (1), 25-38 (1999)
 PUBMED
 10082930
 REFERENCE
 2 (bases 1 to 1753)
 Uhrig, J.F., Prior, A. and Schwenn, J.D.
 Direct Submission
 Submitted (15-JUL-1996) Biochemistry of Plants,
 Ruhr-Universitaet-Bochum, Universitaetsstr. 150, Bochum 44780,
 Germany
 REFERENCE
 3 (bases 1 to 1753)
 Prior, A. and Schwenn, J.D.
 Direct Submission
 Submitted (03-FEB-2000) Biochemistry of Plants,
 Ruhr-Universitaet-Bochum, Bochum 44780, Germany
 JOURNAL
 REMARK
 On Feb 15, 2001 this sequence version replaced gi:1488042.
 COMMENT
 Location/Qualifiers
 1..1753
 /organism="Catharanthus roseus"
 /mol_type="mRNA"
 /db_xref="taxon:4058"
 1..1753
 /gene="par2neu"
 /evidence=experimental
 106..1500
 /gene="par2neu"
 /function="reduction of adenylyl sulfate (APS)"
 /note="C-terminus similar to pdi and beta-subunit of 4-hydroxylase, PDB Accession Number: 1MEK"
 /codon_start=1

Db 1234 GTGTACTTTACGCACCTTGGTCCGCTTTTGGCCAGCAATGGAAGGATCATATCTTGAA 1293

Qy 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416

Db 1294 TTGGCAGAAAAGTTGGCTGTTCT--GGTGAAGAGTTGGAAGTTTAAGGCAGAGGT 1350

Qy 417 GluGlnLysGluPheAlaLysSerGluGlnLeuGlySerPheProThrIleLeuPhe 436

Db 1351 GATCAGAAGGCATTGTCACACAGCAAGATTGCGCTGAATAGCTTCCCTACAAATCTTTTC 1410

Qy 437 PheProLysHisSerSerArgProThrIleLysThrProSerGluLysArgAspValAsp 456

Db 1411 TTCCCAAGCACTCATCTAGCCC--ATAAAATATCCATCAGAGAAGGATGTAGAC 1467

Qy 457 SerLeuMetAlaPheValAsnAlaLeuArg 466

Db 1468 TCATTGATGCTTTTGTATGCTTCCGGA 1497

RESULT 8

CQ805882 1398 bp DNA linear PAT 10-MAY-2004

LOCUS Sequence 2293 from Patent WO2004035798.

DEFINITION CQ805882

ACCESSION CQ805882

VERSION CQ805882.1 GI:47111575

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE 1

AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.

TITLE Identification of novel e2f target genes and use thereof

JOURNAL Patent: WO 2004035798-A 2293 29-APR-2004;

CropDesign N.V. (BE)

FEATURES

source Location/Qualifiers

1. 1398

/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 3.17e-127 Length: 1398

Score: 1811.50 Matches: 351

Percent Similarity: 84.0% Conservative: 48

Best Local Similarity: 73.9% Mismatches: 57

Query Match: 74.1% Indels: 19

DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x CQ805882 (1-1398)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19

Db 1 ATGGCAATGTCGTAAATGTTCTCTCTCTGCTTCGGAATCATAAACTCTCGTTTC 60

Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35

Db 61 GGTGTTTCATGGAGCCAAAGTTTCGCAATTGGTTCGTGAGTTATGGATCGTGT 120

Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54

Db 121 CATGTTGCTCCTGCTCTCTGAATCTATCTGGGAAGCGATCATCATCTGTGTTAAACCTTTA 180

Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74

Db 181 AACGCTGAACCAAGCAAGAAAGGATTCAATGATTCCCTCTGCGGCAACAATGGTAGCAGAA 240

Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

Db 241 ATTGCAGAGGAGTTGAGTGGTGTGAGATTGAGATTGAGATTGAGAGCTTGTAGAGATTGA 300

Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111

Db 301 GAGATGCTTTCACCTCTTGAGATTATGGCAAAAGCTCTTGAGAAATACCGGAACGATATC 360

Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131

Db 361 GCCATTGCAATTAGTGGTGCAGAGATGTGCTCTTATTAGTACGCTCATTTGACTGGG 420

Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151

Db 421 AGGCCATTTAGAGTATTTAGTTTGGATACAGGAGGTTGAATCCTGAGACGTATCGGTTT 480

Qy 152 PheAspAlaValGlnLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171

Db 481 TTCGATGCGGTGAGAGCACTATGGGATTAGGATTAGTATATGTTTCTGATCTCTGTT 540

Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191

Db 541 GAGGTTCAAGGTTTGGTTAGGACAAAGGATGTTCTCTTTTATGAGGATGGTCATCAG 600

Qy 192 GluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAla 211

Db 601 GAGTGTTCGCGTCTTCGAAAGGTGAGACCTTTGAGGCGTCTCTCAAGGGTTTAAAGGCT 660

Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231

Db 661 TGGATTACTGTCAGAGGAAGATCAATCTCCGGGACAAAGGCTCTGAGATTCCGGTTGTT 720

Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251

Db 721 CAGGTTGATCCCGTGTGTTGAGGTTTGGATGTTGGAGTTGTTGTTGGTGAAGTGAAT 780

Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271

Db 781 CCGGTTTGCAGATGTTCAAGGGAATGATGTTTGGAACTCTCTGAGGACTATGGATGTTCCG 840

Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291

Db 841 GTTAACACATTTGCATCGCGCAGGTTATATATCATGATTGGATGTGAGCCTTCACGAAAGCG 900

Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311

Db 901 GTTTTACCGGGTCCAGCAGAGAGAGAGAGTGTGGTGGGAAGATGCTTAAAGCCAG 960

Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331

Db 961 GAATGTGGACTTCACAAAGGAATGTCAAA--GAAACTCCCGATGCTGTAAGTGAAC 1017

Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsnSer 351

Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056

Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371

Db 1057 GAGAACTCTGTGACTTTGAGCAGGAGGAGGATGAGAAATTGATGAAGTTGGAGAACCGT 1116

Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391

Db 1117 AAAGACCTTGGATCGTCTGCTTTATGCTCCGTGGTGGCCCTTTTGTCAAGCCATGAA 1176

Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411

Db 1177 GCATCGTATGTAACATCGCGGATAAATTTGGCTGGAAAGT---GGGATTAAAGTTTGCATA 1233

Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431

Db 1234 TTCAGAGCAGATGGTGCAGAGAGGAGTTGCTTAAGCAGAAATTCGAGCTCGGTAGCTTC 1293

Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451

Db 1294 CCTACCATTTCTGGTTTTCCTAAGAACTCATCGAGACCG---ATCAAGTATCCGTCGAG 1350

Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466

```
Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCGTTCTTGAATCTTGTCCGA 1395
RESULT 9
AX412862 1398 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 626 from Patent WO0222675.
DEFINITION AX412862
ACCESSION AX412862
VERSION AX412862.1 GI:21445320
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.
TITLE Plant genes, the expression of which are altered by pathogen
infection
JOURNAL Patent: WO 0222675-A 626 21-MAR-2002;
Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
Jeffrey L. (US); Eulgem, Thomas (US)
FEATURES
source
1. .1398
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 3.17e-127 Length: 1398
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 6 Gaps: 7
US-10-731-525-8 (1-466) x AX412862 (1-1398)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 1 ATGGCAATGTCTGTAATGTTTCTTCTTCTTCTGCGATCATATAAATCTCGTTTC 60
Qy 20 -----GluProLeuLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 61 GGTGTTTCATTTGGAGCCAAAGATTTCGCAAAATGTTGTTGAGGTTATTCGATCGTGT 120
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 121 CATGTTGCTCTGTGCTCTGAATCTATCTGGGAGCGATCATCATCTGTTAAACCTTTA 180
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 181 AACGCTGAACCAAGACAAAGAGATTCAATGATCTCTTTCGGCAACAATGTTGACGAA 240
Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 241 ATTGCAGAGGAAGTTGAAGTGGTTGAGATTGAGGATTTTGAAGAGCTTCTGAAGAAGTTA 300
Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db 301 GAGAAATGCTTCCCTCTTGAGATTATGCAAAAGCTCTTGAGAAATACGGGACGATATC 360
Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGly 131
Db 361 GCCATTGCAATTTAGTGGTGCAGAGATGTTGCTCTTATTGAGTAGCGTTCATTGACCTGG 420
Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
Db 421 AGGCCATTAGAGATTATTAGTTGATACAGGAGGTTGAATCTCTGAGACGATTCGGTTT 480
Qy 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluThrMetPheProAspAlaVal 171
```

```
481 TTCCATGCGGTGAGAGACACTATGGGATTAGGATTAGATATATGTTTCTCGATTCTGTT 540
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATTTGTTCTCTTTTATGAGGATGTCATCAG 600
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 601 GAGTGTTCGGTGTTCGAAAGGTGAGACCTTTGAGCGCTCTCTCAAGGGGTTAAAGGCT 660
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 661 TGGATTACTGTCAGAGAAAGATCAATCTCCGGGGCAAGGCTCTGAGATTCCGGTTGTT 720
Qy 232 GluValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTyrAsn 251
Db 721 CAGGTTGATCCGGTGTTCGAAAGGTTCGATGGTGGAGTTGTTAGTTTGGTGAAGTGAAT 780
Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 781 CCGGTTGCGAATGTTGAAGGGAATGATGTTTGAACCTTCTTGAGGACTATGGATGTTCCG 840
Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 841 GTTAAACACATTGTCATCGCGCAGGCTATATATCGATTGCGATGTGAGCCTTCACAGAAACG 900
Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311
Db 901 GTTTTACCGGTCAGCACGAGAGAGAGAGATGTTGGTGGGAGAGATGCTTAAAGCCCAAG 960
Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluAspValAsnGlyAsn 331
Db 961 GAATGTGACTTTCACAAAGGGAATGTCAAA--GAAAACTCCGATGATGCTTAAAGTGAAC 1017
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAAGAT 1056
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1057 GAGAAATCTTGATGACTTTGACGAGCGAGGATTTGAGAAATTTGATGAAGTTGGAGAACCGT 1116
Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db 1117 AAAGAGCCTTGGATCGTCGTCGTTTATGCTCCGTTGGTGGCCCTTTGTCAGGCAATGAA 1176
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1177 GCATCGTATGATGAATCGCGGATAAATTCGCTGGAAGT---GGGATTAAGGTTGCCAAA 1233
Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1234 TTCAGACAGATGTTGACCAAGAGAGTTGCTTAAGCAGGAAATTCGAGCTCGTAGCTTC 1293
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1294 CTTACCAATCTCGTTCCTTCCCTAAGAACTCATCGAGACCG--ATCAAGTATCCGCTCAG 1350
Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCGTTCTTGAATCTTGTCCGA 1395
RESULT 10
AX505803 1398 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 498 from Patent WO0216655.
DEFINITION AX505803
ACCESSION AX505803
VERSION AX505803.1 GI:23387040
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
```


Alignment Scores:

Pred. No.: 3,17e-127 Length: 1398
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x AK651565 (1-1398)

```
QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 ATGGCAATGCTCTAAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GGTGTTTCATTTGGAGCCCAAAAGTTTCGCAAAATTTGGTTGATGAGTTATTGGATCGTGT 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 36 IleGlyGlyAlaValAlaPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 CATGTTGCTCTGTGTCTCTGAATCTATCTGGGAAGCGATCATCTCTGTTAAACCTTTTA 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 AsnAlaGluProProAlaGlyLysAspSerIleValProLeuAlaAlaThrIleValAla 74
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 AACGTCGAACCAAGACAAAGGATTCAATGATTCCTCTTGGCGCAACAATGGTAGCAGAA 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 ATTGCAGAGGAAGTTGAATGGTTGAGATTGAGGATTTTGAAGAGCTTGTCTAAGAACTTA 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 GAGAACTGCTCACCTCTTGAGATTATGGAACAAAGCTCTTGAGAAATACGGGAACGATATC 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluIleAlaLysLeuThrGly 131
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 GCCATTGCAATTAGTGGTCGAGAGAGATTGCTCTTATTGAGTACGCTCATTTGACTGGG 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrglnLeu 151
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 AGGCCATTAGATATTAGTTTGGATACAGGAGGTTGAATCTCGACGATATCGGTTT 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 PheAspAlaValGluLysHisTyrglyIleArgIleGluIleMetPheProAspAlaVal 171
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 TTCATGCGGTGGGAAGACACTATGGGATTAGGATTGAGTATATGTTTCTCGATTCGTT 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrgluAspGlyHisGln 191
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATTGTTCTCTTTTATGAGGATGGTCATCAG 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 GAGTGTTCGCCGTGTCGAAGGTCAGACCTTTGAGGCGGTGCTCTCAAGGGGTTAAAGGCT 660
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 TGGATTACTGTCTAGAGAAAGATCAATCTCCGGGCAAGGCTCTGAGATTCCCGTGT 720
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 CAGGTTGATCCGGTGTGTTGAAGGTTTGGATGGTGGAGTTGGTAGTTTGGTGAAGTGAAT 780
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 CCGGTTGCCAATGTTGAAGGAATGATGTTTGGAACTTCTTGGAGACTATGATGATGTTCCG 840
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 272 ValAsnSerLeuHisAlaLysGlyTyrrValSerIleGlyCysGluProCysThrArgPro 291
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 841 GTTAACACATTGTCATCGCGCGAGGGTATATATCGATTGGATGTGAGCCTTGCACCAAGCG 900
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLys 311
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 901 GTTTTACCGGTCAGCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
```

```
QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGlnGluAspValAsnGlyAsn 331
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 961 GAATGTGGACTTTCACAAAGGGAATGCAAA---GAAACTCCGATGATGCTAAAGTGAAC 1017
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1018 GGGGAATCGAATCC-----GCTGTTGCAGATATCTTTAAAGAGT 1056
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1057 GAGAAATCTTGTGACTTTGAGCAGCGAGGGAATGAGAAATTTGATGAAGTTGGAGAACCGT 1116
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 372 LysGluProTrpLeuValValLeuTyrrAlaProTrpCysProTyrrCysGlnAlaMetGlu 391
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1117 AAAGAGCCTTGGATGTCGTCGCTTTATGCTCCGTGGTGGCCCTTTGTGCAAGCCATCGAA 1176
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 GluSerTyrrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1177 GCATCGTATGATGAACCTGGCGGATAAATTCGCTGGAAGT---GGGATTAAGTTGCCAAA 1233
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuLysSerPhe 431
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1234 TTCAGAGCAGATGCTGACCAAGAGGAGTTTGCTTAAGCAGGAATTCGACGCTCGTAGTTC 1293
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrrProSerGlu 451
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1294 CCTACCACTTCGTTTTCCTTAAGACTCATCGAGACCG---ATCAAGTATCCGCTCGAG 1350
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1351 AAGAGAGATCTTGAGTCTTTCGCTTTCGTTCTTGAATCTTGTCCGA 1395
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 12

BT002612 1398 bp mRNA linear PLN 23-DEC-2002
LOCUS Arabidopsis thaliana At4g04610/P4H6_13 mRNA, complete cds.
DEFINITION BT002612
ACCESSION BT002612
VERSION BT002612.1 GI:27363217
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1398)

REFERENCE
AUTHORS Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.

TITLE Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1398)

REFERENCE
AUTHORS Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.

TITLE Direct Submission

JOURNAL

Submited (23-DEC-2002) Salk Institute Genomic Analysis Laboratory
(STGnAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

The discrepancy does not affect the protein sequence.
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shim,D., Bower,L., Chan,M.M., Chang,C.M., Dale,J.M., Heuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

```

source
1..1398
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"
    /db_xref="taxon:3702"
    /chromosome="4"
    /clone="U14453"
    /ecotype="Columbia"
    /note="This clone is in pUNI 51"
1..1398
    /note="5'-adenylyl sulfate reductase"
    /codon_start=1
    /product="At4g04610/P4H6_13"
    /protein_id="AA011528.1"
    /db_xref="GI:27363218"
    /translation="MAMSVNVSSSSGIINSRFSVLSPEKVSQIGSLRLDLRVHAP
VSNLSGRSSVKPLNAEPKTDIMPLAATVIAEVEVEVEIEDPELAKKLE
ASPLEIMDKALEKGNIDIAAFSGAEDVALIEVAHLTGRFVFSLDGTGLNPETVRF
FDVEKHVGRIRIEMPPDSVQGLVRSKGLSPFYEDGHOECRCVRKPLRLAKGL
KAMITGQRKQSGFTRSEIPVQVDPVFEGLDGVSLVKNPVPANVEGNDWNFLRT
MDVPNTLHAAGYISIGCEPCTKAVLPQHERGRWMWEDAKKCEGLHKGNVKENS
DAVNGESKSAVADIFKSENLVLSRQGIENLMKLENRKEPFWLVLYAPWPCQAME
ASYDELADKLAGSGIKVAKPRADQDQKEFAKQBLQLGSFPFTILVFPKNSRRPIKYPSE
KRDESLTSFLNLVR"
misc_difference
507
    /note="compared to genome"
    /replace="g"

ORIGIN
Alignment Scores:
Pred. No.:      3,17e-127      Length:      1398
Score:          1811.50      Matches:      351
Percent Similarity: 84.0%      Conservative: 48
Best Local Similarity: 73.9%      Mismatches:  57
Query Match:     74.1%      Indels:      19
DB:              15          Gaps:         7

US-10-731-525-8 (1-466) x BT002612 (1-1398)
Qy      1  MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerPheProSerSer--- 19
Db      1  ATGCCAATGCTGTAATGTTTCTCTCTCTCGTCTTCGGATCATAAATCTCGTTTC 60
Qy      20  -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db      61  GGTGTTTCATTTGGAGCCAAAGTTTCGCAAAATGGTTTCGTTGAGTTATTGGATCGTGT 120
Qy      36  IleGlyGlyAlaValAlaPheAenLeuSer---GlnArgArgSerLeuValLysProVal 54
Db      121  CATGTTCCTCCTGTGCTCTCGAATCTATCTCGGAAGCGATCATCTGTATAACCTTTA 180
Qy      55  AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
Db      181  AACGCTGAACCAAGACAAAGGATTCAATGATTCCTCTTCGGCGCACAAATGGTAGCAGAA 240
Qy      75  SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db      75  SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

```

```

Db      241  ATTTCAGAGGAAGTTGAAGTGGTTGAGATTGAGGATTTTGAAGAGCTTGTCTAAGAAGTTA 300
Qy      92  AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db      301  GAGNATGCTTCACTCTTTGAGATTATGGACAAGCTCTTTGAGAAATATCGGGAACGATATC 360
Qy      112  AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluIleValAlaLysLeuThrGly 131
Db      361  GCCATTGCAATTAGTGGTGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGG 420
Qy      132  ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluIleThrGlnLeu 151
Db      421  AGGCCATTAGAGATTATTAGTTGGATACAGGAGGTTGAATCTCTGAGACGTATCCGTTT 480
Qy      152  PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
Db      481  TTCGATCGGTGGAGAAGCACTATGGNATTAGNATTGAGTATATGTTTCTGATCTGTT 540
Qy      172  GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db      541  GAGGTTCAAGGTTTGGTTAGGACCAAGGATGTTCTCTTTTATGAGGATGGTCAATCAG 600
Qy      192  GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db      601  GAGTGTTCGGCTGTTCGAAAGGTGAGACCTTTGAGCGCTCTCTCAAGGGTTTAAAGGCT 660
Qy      212  TriPheThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db      661  TGGATTACTGTCAGAGGAAGATCAATCTCCGGGACAAAGGCTTGAGATTCCGGTTGTT 720
Qy      232  GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsn 251
Db      721  CAGGTTGATCCGGTGTGTTGAAGGTTTGGATGGTGGAGTTGTTGTTGGTGAAGTGAAT 780
Qy      252  ProValAlaAsnValLysGlyHisAspIleTyrPheAsnPheLeuArgThrMetAsnValPro 271
Db      781  CCGGTTGCGAATGTTGAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCG 840
Qy      272  ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db      841  GTTAAACATTTGATGGCGGCGGATATATCATTTGGATTGGATGGATGGATGGATGGAT 900
Qy      292  ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpGluAspAlaLysAlaLys 311
Db      901  GTTTTACCGGTCAGCACGAGAGAAGGAGATGGTGGTGGGAGAGATGCTTAAAGCAAG 960
Qy      312  GluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsn 331
Db      961  GAATGTGGACTTCACAAAGGGAATGTCAA--GAAAACTCCGATGCTTAAAGTGAAC 1017
Qy      332  GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db      1018  GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056
Qy      352  ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db      1057  GAGAACTCTTGATCTTTGAGCAGCAGGAGGATTTGAGAAATTTGATGAAGTTGGAGAACCGT 1116
Qy      372  LysGluProThrLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGlu 391
Db      1117  AAAGAGCCTTGGATCGTCTGCTGCTTTATGCTCGGTGGTGGCCCTTTTGTCAAGCATGGAA 1176
Qy      392  GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db      1177  GCATCGTATGATGATGCGGATTAATTTGGCTGGGAAGT---GGGATTAAAGTTGCCAA 1233
Qy      412  PheArgAlaAspGlyGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db      1234  TTCAGAGCAGATGTCAGCAGAGGAGTTGCTTAAGCAGGAATTCGAGCTCGTAGCTTC 1293
Qy      432  ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db      1294  CCTACCATTTCTGGTTTTTCCCTAAGAACTCATCGAGACCG---ATCAAGTATCGCTCTGAG 1350

```



```

Qy 452 LyAArgaspValaaspSerLeuMetAlaPheValaasnAlalaLeuArg 466
Db 1351 AAGAGAGATGTTGAGTCTTGACTTCGTTCTTCTGAATCTTGTCGA 1395

RESULT 13
AF424582 1710 bp mRNA linear PLN 08-OCT-2001
LOCUS Arabidopsis thaliana AT4g04610/F4H6_13 mRNA, complete cds.
DEFINITION
ACCESSION AF424582
VERSION AF424582.1 GI:15983415
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1710)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1710)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (26-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
Arabidopsis Full-Length cDNA) : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
Location/Qualifiers
1..1710
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="RAFL09-34-D13(R14453)"
/ecotype="Columbia"
1..63
64..1461
/note="5'-adenylylsulfate reductase"
/codon_start=1
/product="AT4g04610/F4H6_13"

FEATURES
source
5'UTR
CDS

```

```

/protein_id="AAL11576.1"
/db_xref="GI:15983416"
/translation="MAMSVNVSSSSSGIINSRFGVSLRPKVSIQSLRLDRVHVAP
VSLNLSKRSVSKPLNAPKTKDMSPIAATWVAETAEVEVEVEIEDEFEELAKLEN
ASPLEIMKALEKYGNDIAIAFSGAEDVALIEYAHILTRFRFRVSLDTRLNRPETYRF
DFAVEKHGIRIEYMPDSVEVQGLVRSKGLFSEYEDHQECRCRVRPLRRALKGL
MVTGQRKDQSPGTRSEIPVQVDVPFEGLDGVSGLVKWNPVANVGVNDVWNLFLT
MDVPVNTLHAGYISIGCPCTKAVLPQGOHERGRWMDAKAKECCGLHKGNVKNSD
DAKNGESKAVADIFKSNLVTLSRQGIENLMKLENRKEPWIWLYAPWCPQOME
ASYDELALAGSGIKVAKPRADGQKQFAKQLQLGQSFPIILVFPKNSRPIKYPSE
KRDVESLTSPFLNLR"
1462..1710
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 4,02e-127 Length: 1710
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 15 Gaps: 7
US-10-731-525-8 (1-466) x AF424582 (1-1710)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 64 ATGCAATGTCGTAAATGTTCTTCTCTCTGCTTCTCTGGGATCATAAACTCGTTTC 123
Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 124 GGTGTTTCATTTGGAGCCAAAGTTTCGCAATTTGTTCTGTTGAGTTATTGGATCGTGT 183
Qy 36 IleGlyGlyAlaValAasnPheAasnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 184 CATGTTGCTCTGCTGTCTCTGAATCTATCTGGGAGCGCATCATCATCTGTTAAACCTTTA 243
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 244 AACCTGACCAACAAAGAGATTCAATGATTCCTCTTGGGGCAACAAATGGTAGCAGAA 303
Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 304 ATTGCAGAGGAAGTTGAAGTGTGTTGAGATTGAGGATTTTGAAGAGCTTGTCTAAGAAGTTA 363
Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db 364 GAGAATGCTTCACCTCTTGAGATTATGGACAAAGCTCTTTGAGAAATACGGGAACGATATC 423
Qy 112 AlaIleAlaPheSerGlyValAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db 424 GCCATTGCATTTAGTGTGCGAAGATGTTGCTCTTATTAGTAGTACGCTCATTTGACTGGG 483
Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAasnProGluThrTyrGlnLeu 151
Db 484 AGGCCATTAGAGTATTAGTTTGGATACAGGAGGTTGAATCTCTGAGACGATATCGGTTT 543
Qy 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
Db 544 TTCGATGCGGTGGGAAGACACTATGGGATTAGGATTGATATATGTTTCTCGATTCTGTT 603
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 604 GAGGTTCAAGTTTGGTTAGGAGCAGGATGTTCTCTTTTATGAGGATGGTCAATCAG 663
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 664 GAGTGTTCGCGTGTTCGAAAGCTGAGACCTTTGAGCGCTGCTCTCAAGGGTTTAAAGGCT 723
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 724 TGGATTACTGTCAGAGGAAAGATCAATCTCCGGGGAACAAGCTCTGAGATTCGGGTGTT 783
Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251

```



```
DB: 15 Gaps: 7
US-10-731-525-8 (1-466) x AY088011 (1-1714)
QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrThrThrPheProSerSer--- 19
Db 67 ATGGCAATGCTGTAATGTTCTTCTCTCTGCTCTCTGGGATCATAAACTCTCGTTTC 126
QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluAtrPro 35
Db 127 GGTGTTTCATTTGGAGCCAAAGTTTCGCAATTGTTGCTTGGATTATTGGATCGGT 186
QY 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 187 CATGTTGCTCCTGCTGCTCTGCAATCTATCTGGGAAGCGATCATCTGTTAAACCTTTA 246
QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
Db 247 AACCTGACCAAGACAAACGATTCAATGATTCCTCTTGGCGCAACAATGGTAGCGAA 306
QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 307 ATTGCAGAGGAGTTGAAGTGTGAGATTGAGGATTTTGAAGCTTGTCTAAGAGTTA 366
QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db 367 GAGAAATGCTTTCACCTCTTGAGATTATGACAAAGCTCTTGAGAAATACCGGGAACGATATC 426
QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db 427 GCATTCGATTTAGTGTGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGG 486
QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
Db 487 AGCCATTAGATATTAGTTTGGATACAGGAGGTTGATCTCGAGAGTATCGGTTT 546
QY 152 PheAspAlaValGlnLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
Db 547 TTCGATCGGTGGAGAGCACATATGGGATTAGGATTGAGTATATGTTTCTCGATTCTGTT 606
QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 607 GAGTTCCAGGTTTGGTTAGGACAGGAGATTGTTCTCTTTTATGAGGATGTCATCAG 666
QY 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 667 GAGTTGTCGCTGTCGAAAGGTGAGACCTTTGAGCGTGCTCTCAAGGGTTTAAGGCT 726
QY 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 727 TGGATTACTGCTCAGAGGAAAGATCAATCTCCGGGACAAAGTCTGAGATTCCGGTTGTT 786
QY 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251
Db 787 CAGGTTGATCCGGTGTGAAAGTTTGGATGGTGGAGTGGTGTGTTGGTGAAGTGGAAAT 846
QY 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 847 CCGGTTGCGAATGTTGAGGGAATGATGTTTGGAACTTCTTGGAGCATGAGATGTTCCG 906
QY 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 907 GTTAAACACATTGTCATCGCGGCGGTATATATCGATTGGATGTGAGCTTGCACGAAAGCG 966
QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLys 311
Db 967 GTTTTACCCTGGTTCAGCACGAGAGGAGATGTTGGTGGGAAATGCTTAAAGCCAAG 1026
QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsn 331
Db 1027 GATGTGGACTTCACAAAGGAATGTCAAA--GAAACCTCCGATGCTAAAGTGAAC 1083
QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
```

Search completed: February 21, 2006, 00:18:49
Job time : 7271 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 20, 2006, 21:57:15 ; Search time 840 Seconds

(without alignments)
3697.318 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPTSTTPSSSE.....KYPSEKRDVDSLMAFYNALR 466

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Ygapop 10.0			
Fgapop 6.0			
Delop 6.0			

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlh
-Q=/abs/ABSSWEB_spool/US10731525/runat_17022006_164855_69/app_query.fasta.1
-DB=N_geneseq -QMT=fascap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USPR=US10731525 @CGN 1 1 727 @runat_17022006_164855_69 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

No.	Score	Query	Length	ID	Description
1:	Geneseqn1980s.*				
2:	Geneseqn1990s.*				
3:	Geneseqn2000s.*				
4:	Geneseqn2001as.*				
5:	Geneseqn2001bs.*				
6:	Geneseqn2002as.*				
7:	Geneseqn2002bs.*				
8:	Geneseqn2003as.*				
9:	Geneseqn2003bs.*				
10:	Geneseqn2003cs.*				
11:	Geneseqn2003ds.*				
12:	Geneseqn2004as.*				
13:	Geneseqn2004bs.*				
14:	Geneseqn2005s.*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	2445	100.0	1629	3	AZ50066 Soybean a
2	1951	79.8	1730	13	ADX13202 Plant ful
3	1923	78.7	1795	3	AZ50065 Soybean a
4	1811.5	74.1	1398	6	ABZ12693 Arabidops

5	1811.5	74.1	1398	6	ADG88184	Adg88184 A. thalia
6	1811.5	74.1	1398	8	ADA68112	Ada68112 Arabidops
7	1811.5	74.1	1398	12	ADN74398	Adn74398 Thale cre
8	1811.5	74.1	1712	3	AAC47286	Aac47286 Arabidops
9	1806.5	73.9	1714	3	AAC37691	Aac37691 Arabidops
10	1801	73.7	1702	3	AAC38867	Aac38867 Arabidops
11	1797	73.5	1377	3	AAC34415	Aac34415 Arabidops
12	1758.5	71.9	1687	3	AAC32729	Aac32729 Arabidops
13	1745.5	71.4	110000	3	AAP22305_03	Abt110000 (4 of
14	1738	71.1	1365	6	ABZ14500	Abz14500 Arabidops
15	1738	71.1	1365	6	ADG88214	Adg88214 A. thalia
16	1738	71.1	1365	6	ADG88215	Adg88215 A. thalia
17	1738	71.1	1365	12	ADN73974	Adn73974 Thale cre
18	1738	71.1	1685	3	AAC48356	Aac48356 Arabidops
19	1645.5	67.3	1827	3	AZ50067	Az50067 Wheat ade
20	1645.5	67.3	1840	13	ADT17458	Adt17458 Plant CDN
21	1597.5	65.3	1380	4	AAH42329	Aah42329 Nucleotid
22	1295.5	53.0	1210	3	AZ50064	Aaz50064 Impatiens
23	1262.5	51.6	1215	3	AZ50063	Aaz50063 Corn aden
24	1241.5	50.8	1183	13	ADX32480	Adx32480 Plant ful
25	1203.5	49.2	1156	13	ADX29365	Adx29365 Plant ful
26	862	35.3	1031	3	AACS2097	Aac52097 Arabidops
C 27	860	35.2	588	9	ACL21376	Acl21376 DNA clone
C 28	856	35.0	670	9	ACL21371	Acl21371 DNA clone
C 29	811.5	33.2	783	9	ADA29776	Ada29776 DNA encod
C 30	757.5	31.0	554	9	ACL21377	Acl21377 DNA clone
C 31	745.5	30.5	810	11	ABD15340	Abd15340 Pseudomon
C 32	731.5	29.9	705	13	ADS61808	Ad61808 Bacterial
C 33	716.5	29.3	471	12	ADP93046	Adp93046 Cotton ex
C 34	697.5	28.5	585	13	ACN49050	Acn49050 Cotton pr
C 35	697	28.5	422	9	ACL21373	Acl21373 DNA clone
C 36	679	27.8	623	13	ACN56923	Acn56923 Cotton gy
C 37	646	26.4	475	9	ACL21375	Acl21375 DNA clone
C 38	623	25.5	659	9	ACL21370	Acl21370 DNA clone
C 39	587	24.0	441	9	ACL21374	Acl21374 DNA clone
C 40	546.5	22.4	6063	14	ACL64344	Acl64344 M. xanthu
C 41	540	22.1	723	14	ACL171515	Acl171515 M. xanthu
C 42	539	22.0	690	13	ADT44456	Adt44456 Bacterial
C 43	525	21.5	700	9	ACL21379	Acl21379 DNA clone
C 44	520	21.3	588	13	ACN49122	Acn49122 Cotton pr
C 45	509	20.8	552	13	ACN56966	Acn56966 Cotton gy

ALIGNMENTS

RESULT 1
AAZ50066
ID AAZ50066 standard; cDNA; 1629 BP.
XX
AC AAZ50066;
XX
DT 04-MAY-2000 (first entry)
XX
DE Soybean adenosine-5'-phosphosulfate reductase cDNA clone s12.pk0064.g4.
XX
KW Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;
KW sulphate assimilation protein; SAP; clone s12.pk0064.g4; ss.
XX
OS Glycine max.
XX
FH Key
CDS 32..1432
FT /tag= a
FT /product= "Soybean APS reductase"
XX
PN WO200004161-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015808.
XX
PR 14-JUL-1998; 98US-0092833P.
XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Falco SC, Allen SM;
 XX WPI; 2000-182427/16.
 DR P-PSDB; AAZ50066.
 XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
 PT and seeds useful as probes for isolating cDNAs and genes encoding
 PT homologous proteins, and in producing transgenic plants.
 XX Claim 3; Page 34-35; 47pp; English.
 XX The present cDNA sequence encodes soybean adenosine-5'-phosphosulfate
 CC (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is
 CC designated 812.pk064.g4, isolated from soybean two week old developing
 CC seedlings treated with 2.5ppm chlorimuron. APS reductase polypeptides are
 CC useful for producing antibodies for screening cDNA expression libraries
 CC to isolate full length APS reductase cDNA clones. Chimeric genes
 CC comprising the APS reductase gene when transformed in a host cell can
 CC alter the level of SAP expression. This can be used to increase sulphur
 CC containing compounds in the cell, including the nutritionally critical
 CC amino acids cysteine and methionine
 XX SQ Sequence 1629 BP; 461 A; 309 C; 420 G; 439 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,16e-226 Length: 1629
 Score: 2445.00 Matches: 466
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-731-525-8 (1-466) x AAZ50066 (1-1629)
 Qy 1 MetAlaLeuAlaPheThrSerSerLeSerAlaPheThrSerThrPheProSerSerGlu 20
 Db 32 ATGGCCCTCGCTTTCACCTTCTCAATTTCCGACCAATCCACCTTCCCATCATCGAA 91
 Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyAlaVal 40
 Db 92 CCCAACTTCGCAAAATGGGTCAATTAGGATTTCCGAGAGCCCATTTGGAGCGCGTT 151
 Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
 Db 152 AATTTCAATTTATCTCAAGACGGAGCTTGGTAAAGCCGTTAAACGCCGAACCTCCACGC 211
 Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys 80
 Db 212 AAGGATTCATGTGTTCTCTCGACGCAACACCATCGTTGCTTCTCTGACACGAA 271
 Qy 81 GluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100
 Db 272 GAGGAGATTTGAACAGATAGCCAGTCATCTCGACATGCTTCCACCTCTTGAATCATG 331
 Qy 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120
 Db 332 GATAGAGCCCTCGCAAAATTCGCAACGACATAGCTATTGCTTTCAGTGGTGTGAGAT 391
 Qy 121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140
 Db 392 GTTGCTTTGATGATGATGCGAAATGACGGGTGCGACCTTTTAGGGTTTTCAGTTTGGAC 451
 Qy 141 ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly 160
 Db 452 ACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTGAGAGCATTTATGGA 511
 Qy 161 IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys 180
 Db 512 ATTCGATTTAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCAATTTGGTGGAGTAA 571
 Qy 181 GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg 200
 Db 572 GGGTTATTTCTTTCTACGAGGATGGGCACCAAGAGTGTTCAGGGTGACAAAGGTGAGG 631
 Qy 201 ProLeuArgArgAlaLeuLysGlyLeuArgAlaTyrIleThrGlyGlnArgLysAspGln 220
 Db 632 CCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGATAAATGGTTCAGAGAAAGACAG 691
 Qy 221 SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet 240
 Db 692 TCACCTGGTACTAGGCTGAAATACCGGTGTTTCAGGTTGATCCGGCTTTTGGAGGAATG 751
 Qy 241 AspGlyGlyIleGlySerLeuValLysThrAsnProValAlaAlaAsnValLysGlyHisAsp 260
 Db 752 GATGTTGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGGCCATGAC 811
 Qy 261 IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr 280
 Db 812 ATATGGAACTTCTTAGGACCATGAATGTCCTGTGAATTCCTTGCATGCAAAAGGATAT 871
 Qy 281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu 300
 Db 872 GTTTCCAAATGGGTGTGAGCCCTGCACATAGGCTGTTTACCTGGGCAACATGAAGGGAA 931
 Qy 301 GlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320
 Db 932 GGGAGGTGGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGCTTCTTCAAAAGGAATGTA 991
 Qy 321 LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly 340
 Db 992 AAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGGT 1051
 Qy 341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr 360
 Db 1052 GATGCTACCACTGTGCTCTGACATTTTCAACAGCCGCAATGTAGTTAATCTTGAGCAGACT 1111
 Qy 361 GlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyr 380
 Db 1112 GGAATTCAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGTCTTTAT 1171
 Qy 381 AlaProTrpCysProTyrCysGlnAlaMetGluLysSerTyrValAspLeuAlaAspLys 400
 Db 1172 GCACCATGGTGGCCCTACTGCCAGGCTATGAGGAATCTTTATGTGTGACTTAGCAGACAG 1231
 Qy 401 LeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGlu 420
 Db 1232 TTACAGGCTCAACAGGATGAAGTTGGAATAATTAGACGATGGAGAACAGAAAGAA 1291
 Qy 421 PheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHis 440
 Db 1292 TTTGCAAAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTATTTTCCCAAGCAT 1351
 Qy 441 SerSerArgProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAla 460
 Db 1352 TCGTCTCGGCCCAACAATAAAGTATCCCTCGAAAGAGAGATGTTGATTCTCTTGATGCA 1411
 Qy 461 PheValAsnAlaLeuArg 466
 Db 1412 TTTGTAATGCTTTAAGA 1429
 RESULT 2
 ADX13202
 ID ADX13202 standard; cDNA; 1730 BP.
 XX
 AC ADX13202;
 XX 21-APR-2005 (first entry)
 DT
 XX Plant full length insert polynucleotide seqid 7777.
 DE
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;

growth rate; cell cycle pathway; disease resistance;
galactomannan production; lignin production; plant growth regulator;
yield; plant growth; plant development; seed oil; protein yield;
protein content; gene; ss.

Unidentified.

US2004034888-A1.

19-FEB-2004.

28-APR-2003: 2003US-00425114.

06-MAY-1999: 99US-00304517.

8/9C8600-S0T007 ; T007-ACN-C0

(LNUJ/) LIU J.

(KOVA/) KOVALI
(SCPE/) SCPEFI

(TABA//) TABASKA J E.
(CAOV//) CAO V

THE T. THOMAS Y.

WDT: 2004 100133/15

— 2 —

cold, heat, drought, herbicides, extreme osmotic conditions,

improving yield.

Claim 1; SEQ ID NO 7777; 15pp; English.

The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at

of the invention are also useful in physical arrays of molecules and

improving plant

osmotic conditions, pathogens or pests, plant cells by modification of the cell

increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This sequence represents a plant full length insert polynucleotide that can be used in the recombinant DNA construct of the invention.

Sequence 1730 BP: 448 A; 350 C; 427 G; 505 T; 0 U; 0 Other:

Assignment Scores:

Length:	1951.00	Matches:	38
Conservative:	88.2%	Conservative:	38
Local Similarity:	80.3%	Mismatches:	40
/ Match:	79.8%	Indels:	16
	13	Gaps:	7

0-731-525-8 (1-466) x ADX13202 (1-1730)

1 MetAlaLeuAlaPhe-----ThrSerIleSerAlaProThrSerThr 15
40 ATGCTCTCGCGTTTCCACTACTTCTCTTCAGCAGCTGTGCAGCTTCAGTCTCC 99

16 Phe-----ProSerSerGluProlLysLeuProGlnIleGlySerIleArg 30
||| |||||:: ||| ||||| ||||| ||||| ||||| |||||
100 TTCTTCTCTCGTCTCGATCTTCTTCGGACGCCAAAGCTCCTCAAATCGGTTCGTTTCGG 150

Qy	31	IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnIleArgSerLeu	50
Db	160	TTTCCGGAGAGGTCGCTTGTCTGTTGTTGTTCAATGTAATCAACGACGCTCCTTG	219
Qy	51	ValIysProValAsnAlaGluProProArgTlyAspSerIleValProLeuAlaAlaThr	70
Db	220	GTGAGGCCATCTCAACGCCGAACCGAACGAAATTCATTGTTCTCTTCGAGCA---	276
Qy	71	ThrIleValAlaSerAlaSerGluThrIleGluGluAspPheGluGlnIleAlaSerAsp	90
Db	277	ACTATCGTCTCTCTGAGGTTTGAGAGGAGAGGAGATTTTGAGCNAATAGCGAAGAC	336
Qy	91	LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspIysPheGlyAsnAsp	110
Db	337	CTTGAAATGCATCTCTCTTTGAAATATATGATAGGCCCTCGAGAAATTTGGCAACGAC	396
Qy	111	IleAlaIleAlaPheSerGlyAlaGluAspValAlaIleuIleGluTyrAlaIysLeuThr	130
Db	397	ATCGCTATTGCATTTATGTGTCTGTAAGATGTTCCTTTGATTGAGTATGCACATTTGACG	456
Qy	131	GlyArgProPheArgValPheSerLeuAspThrGlyIleValLeuAsnProGluThrTyrGln	150
Db	457	GGTCGACCTTACAGAGTTTATGCTTTGACACATGGGAGACTGTAATCCGAAACCTACAA	516
Qy	151	LeuPheAspAlaValGluIysHisTyrGlyIleArgIleGluTyrMetPheProAspAla	170
Db	517	TTTTTTTGATCGGTTGAGAGCAATATGGAATTCACATTCAGTATCATGTTCCCTGATCG	576
Qy	171	ValGluValGlnAlaLeuValArgSerLysGlyIleuPheSerPheTyrGluAspGlyHis	190
Db	577	GTTGAGGTTTCAGGCATTTAGTAAGAACTAAAGGCTCTTCTCATTTTACGAGGATGGGCAT	636
Qy	191	GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuIysGlyLeuArg	210
Db	637	CAGAAGTGTCCGATGAGAGAGGTGAGGCCCTTCGAGGAGACCCCTTAAGGGTCTCAA	696
Qy	211	AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal	230
Db	697	GCATGGATCACTGGACAGAGAAAGATCACTCTCTGGTACTAGGTCTGAAATCCCTGTT	756
Qy	231	ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValIysTrp	250
Db	757	GTCACGGTTGATCCTGTGTTTGTAGGACACTGGATGGTGGAAATGGCAGCCTGGTGAAGTGG	816
Qy	251	AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal	270
Db	817	AACCCGGTTGCAANTGTTAATGTGCTAGACATATGAGGCTTCTTAGGACCATGGAGATT	876
Qy	271	ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg	290
Db	877	CCTGTAATTCATTGCATTCCCAAGGATATGTTTCGATTGGCTGTGAGCCGTGCACAAGG	936
Qy	291	ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla	310
Db	937	CCAGTTTATCCGGGTCAACATGAAAGAGAGGAAGGTGGTGGTGGGAGGATGCCAAGCC	996
Qy	311	LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnsGluGluAspValAsnGly	330
Db	997	AAGGAGTGTGTCTTTCACAAAGGTATATATAAACATGAGAGATGCTGCCACGCTTAATGGA	1056
Qy	331	AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn	350
Db	1057	AATGGGGCCTCCCAA-----GCAATGGATCTGCCC---ACTGTCGCTGACATTTTCAAC	1107
Qy	351	SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp	370
Db	1108	TCCCAGGATGTGCTGACGTTGACAGGTCTGGGATTCAGAAATTTGGCAAAATTTGGAGAAC	1167
Qy	371	ArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet	390
Db	1168	CGAAAAAGAACCCCTGGCTTGTGTGCTCTATGCAACCATGGTGGCCGCTTCTGTCTAGGCTATG	1227
Qy	391	GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly	410

```

Db      1228 GAGGAATCATATGTTGATTTGGCAGAAAGTTAGCTGGGTCA---GGAGTGAAGTTGCA 1284
Qy      411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db      1285 AAATTCAGCGGATGAGACAGACAGAGGAATATGCAAAAGACTGAATGCACTGGAGTGGGAAGC 1344
Qy      431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db      1345 TTCCCCCAATACTACTCTTCCCAACACACTCTCTCAACCA---ATTAAAGTACCCCTTCG 1401
Qy      451 GluLysArgAspValAspSerLeuMetAlaPheValAlaPheValAlaLeuArg 466
Db      1402 GAAAGAGAGATGTTGATTCATTGACCGCATTCGTGAATGCTTTACGG 1449

```

RESULT 3

AAZ50065
ID AAZ50065 standard; cDNA; 1795 BP.

XX AC AAZ50065;

XX 04-MAY-2000 (first entry)

XX Soybean adenosine-5'-phosphosulfate reductase cDNA clone se4.11g09.

XX Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;
XX sulphate assimilation protein; SAP; clone se4.11g09; ss.

XX Glycine max.

XX Key Location/Qualifiers

XX CDS 50..1468

XX /*tag= a

XX /product= "Soybean APS reductase"

XX WO200004161-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015808.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Allen SM;

XX WPI; 2000-182427/16.

XX P-PSDB; AAY44744.

XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX and seeds useful as probes for isolating cDNAs and genes encoding
XX homologous proteins, and in producing transgenic plants.

XX Claim 3; Page 32-33; 47pp; English.

XX The present cDNA sequence encodes soybean adenosine-5'-phosphosulfate
XX (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is
XX designated se4.11g09, isolated from soybean embryo, 19 days after
XX flowering. APS reductase polypeptides are useful for producing antibodies
XX for screening cDNA expression libraries to isolate full length APS
XX reductase cDNA clones. Chimeric genes comprising the APS reductase gene
XX when transformed in a host cell can alter the level of SAP expression.
XX This can be used to increase sulphur containing compounds in the cell,
XX including the nutritionally critical amino acids cysteine and methionine

XX Sequence 1795 BP; 504 A; 351 C; 436 G; 504 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.3e-176 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservative: 37
Best Local Similarity: 79.4% Mismatches: 45

```

Query Match: 78.7% Indels: 16
DB: 3 Gaps: 7

US-10-731-525-8 (1-466) x AAZ50065 (1-1795)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 56
56 CTTGCCGTTTCCACTACTTCTCTCTCTCAGCTGCAGCAGCAGCAGCGTCTCGAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 116
116 TTCTTCTCGCGCCTTGGATCTTCATCGACGCTAAAGCTCCGCAAAATTTGTTCTTTCGG 175
Qy 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeu 50
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 176
176 TTCCCGGAGAGGCTCAAGTTTCGTCGGTGTGTTTAATTTAACTCAAGACGCTCTCTCG 235
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 236
236 GTGAGGCCACTCAATGCCGAACCGCAACGGAATGATTCTGTGTTCTCTTGCAGCA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 293
293 ACTATCGTTCTCTCTGAGTTGAGAGGAGAAAGATTTTGAGCAATTAGCCAAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 353
353 CTTGAAATTCATCTCTCTTGAGATTATGATTAAGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 413
413 ATCGTATTGCTTGTAGTGTGCTGAAGATGTTGTTGTTGATTGATGACATTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 473
473 GGTGACCTTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAACCCAGAACCTACAAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 533
533 TTTTGTGACGCTGTTGAGAGCATATTGGAATTCATATTGATACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 593
593 GTTGAGGTTTCAGGCATTAGTAAGAACTAAGGGGCTCTCTCATTTTACGAGGATGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 653
653 CAAGAGTGCTGTAGAGTAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGGTCTCAA 712
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 713
713 GCATGGATTACTGGACAGAGAAAGACAGCTCTCTGTTACTAGGTCTGAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 773
773 GTCCAGGTTGATCTGTTTTTTGAGGGAGCTGATGGTGAATTTGGCAGCCTGCTGAAGTGG 832
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 833
833 AACCCGGTGTCAATGTTAATGGTCTAGACATATGGAACTTCCTTAGGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 893
893 CCTGTAAATTCATTGCAATCCCAAGGATATGTTTCGATTTGGCTGTGACCATGCACAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGluArgTrpTrpTrpGluAspAlaLysAla 310
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 953
953 CCGGTTTTTACCCGACCAACATGAAAGAGAGAAAGGTGGTGGTGGAGAGTGCCTGAAGGCC 1012
Qy 311 LysGluCysGlyLeuHisLysLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1013
1013 AAGGAGTGTGCTCTTCAAAAGGTAATTTTGAACAGGAAGATGTCGCCACGCTTAATGGA 1072

```


Qy	331	AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn	350
Db	1073	AATGGGACCTCCCAA-----GGAAATGGCTCTGCC---ACTGTTGCTGACATTTTCATC	1123
Qy	351	SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp	370
Db	1124	TCCCAGAAATGGTTCAGCTTGAGCAGGTCCGGGATTGAGAAATTTGGCAAAATTAGAGAAC	1183
Qy	371	ArgLysGluProThrLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet	390
Db	1184	CGAAAAGAACACTGGCTGTGTGCTCTATGACCATGGTGCCTCTCTCAGGCTATG	1243
Qy	391	GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly	410
Db	1244	GAGAGTCGTATGTGTGATCTGGCAGAGAGTTAGCAAGGTCA---GGAGTGAAGGTTGCA	1300
Qy	411	LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer	430
Db	1301	AAATTCAGAGCCGATGAGAGCAGAGAGATATGCAAGAGTGAATCGAGTTGGGAAGC	1360
Qy	431	PheProThrLeuLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer	450
Db	1361	TTCCCCACAATACTCTCTCTCCCCAAGCACTCTTCTCAACCA---ATTAAGTACCCTTCA	1417
Qy	451	GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg	466
Db	1418	GAAGAAGAGAGATGTTGATTCTATTGACGGCATTCGTGAATGCCTTACGG	1465
RESULT	4		
ID	ABZ12693		
XX	ABZ12693	standard; DNA; 1398 BP.	
AC	ABZ12693;		
DT	21-JAN-2003	(first entry)	
XX			
DE	Arabidopsis thaliana	stress regulated gene SEQ ID NO 498.	
KW	Arabidopsis thaliana;	plant; gene; stress; transgenic; ds.	
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	WO200216655-A2.		
XX			
PD	28-FEB-2002.		
XX			
PF	24-AUG-2001; 2001WO-US026685.		
XX			
PR	24-AUG-2000; 2000US-0227866P.		
PR	26-JAN-2001; 2001US-0264647P.		
PR	22-JUN-2001; 2001US-0300111P.		
XX			
XX	(SCRI) SCRIPPS RES INST.		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Harper JF, Krebs J, Wang X, Zhu T;		
XX			
DR	WPI; 2002-304127/34.		
XX			
PT	Identifying a stress condition to which a plant cell has been exposed and		
PT	producing plants with increased tolerance to these abiotic stresses.		
XX			
PS	Claim 14; SEQ ID NO 498; 577pp + Sequence Listing; English.		
XX			
CC	The invention relates to identifying a stress condition to which a plant		
CC	cell has been exposed, comprising: (a) contacting nucleic acid		
CC	representative of expressed polynucleotides in the plant cell with an		
CC	array or probes representative of the plant cell genome; and (b)		
CC	detecting a profile of expressed polynucleotides in the plant cell		
CC	characteristic of a stress response. The method is useful in the		
CC	production of transgenic plants, cells and seeds and in producing plants		
CC	with increased tolerance to abiotic stress. The present sequence is that		
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used		

```
Db      841  GTTAACACATGTCATCGCGCAGGGTATATATCATGATGTGAGCCCTTCACGAAAGCG 900
Qy      292  ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLys 311
Db      901  GTTTTACCGGTGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy      312  GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331
Db      961  GAATGTGGACTTCACAAGGAATGTCAAA---GAAACTCCGATGATGCTTAAGTCAAC 1017
Qy      332  GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspAllePheAsnSer 351
Db      1018  GGGGAATCGAATCC-----CCTGTGCAGATATCTTTAAGAGT 1056
Qy      352  ProAsnValValAsnLeuSerArgThrGlyLeuGlnAsnLeuAlaLysLeuGluAspArg 371
Db      1057  GAGAATCTTGACTTTGAGCAGCGAGCGGATTTGAGAAATTTGATGAAGTTGGAGAACCGT 1116
Qy      372  LysGluProTrpLeuValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db      1117  AAAGAGCCTTGGATCGTGTCTTATGCTCCGGTGGTCCCTTTTGTCAAGCCATCGAA 1176
Qy      392  GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db      1177  GCATCGTATGATGAATGCGCGGATAAATTGCTGGAAGT---GGGATTAAGTTGCCAAA 1233
Qy      412  PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db      1234  TTCAGACAGATGTCGACCAAGAGAGTTTCTAAGCAGGAATTCGAGCTCGGTAGCTTC 1293
Qy      432  ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db      1294  CCTACCAATTCGTGTTTCCCTTAAAGAACTCATCGAGACCG---ATCAAGTATCCGCTCGAG 1350
Qy      452  LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db      1351  AAGAGAGATGTTGAGTCTTTGACTTCGTTCTTGAATCTTGTCGA 1395
RESULT 5
ADG88184
ID      ADG88184 standard; cDNA; 1398 BP.
XX      AC      ADG88184;
XX      DT      22-APR-2004 (first entry)
XX      DE      A. thaliana RPP4-upregulated pathogen infection-related gene #626.
XX      KW      pathogen infection-related gene; plant; Peronospora parasitica;
XX      KW      defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
XX      KW      fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX      OS      Arabidopsis thaliana.
XX      EN      WO200222675-A2.
XX      PD      21-MAR-2002.
XX      PF      14-SEP-2001; 2001WO-US028506.
XX      PR      15-SEP-2000; 2000US-0232778P.
XX      PR      22-JUN-2001; 2001US-0300183P.
XX      PR      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX      PA      (UYNC-) UNIV NORTH CAROLINA.
XX      PA      (GLAZ/) GLAZEBROOK J.
XX      PA      (WANG/) WANG X.
XX      PA      (DANGL) DANGL J L.
XX      PA      (EULG) EULGEM T.
XX      PA      (ZHUT/) ZHU T.
XX      PI      Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
```

```
XX      WP1; 2002-292409/33.
XX      DR      Novel isolated polynucleotide, useful for conveying pathogen resistance
XX      PT      to plants, and for identifying plants infected with a pathogen.
XX      PS      Claim 3; SEQ ID NO 626; 605pp; English.
XX      CC      The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC      ADG87557)) whose expression is altered in response to pathogen infection,
CC      and to homologues of these genes from other plants or fungi, especially
CC      from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC      cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC      expression of genes of the invention was upregulated or downregulated in
CC      Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC      indicating that they play a role in defence mechanisms. The genes of the
CC      invention are regulated by RPP7 or RPP8 which act via unconventional
CC      signalling cascades, or by the RPP4-dependent pathway. The invention also
CC      relates to polypeptides encoded by the pathogen infection-related genes;
CC      promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC      ; expression cassettes, host cells and pathogen-resistant transgenic
CC      plants and their progeny comprising a polynucleotide of the invention;
CC      and a method of identifying a plant cell infected with a pathogen. The
CC      polynucleotide sequences and methods of the invention are useful for
CC      identifying plants infected with a pathogen, and for conferring
CC      resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC      nematodes and insects (e.g., aphids). The present sequence represents an
CC      Arabidopsis thaliana gene whose expression is altered in response to
CC      Peronospora parasitica infection. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences.
XX      SQ      Sequence 1398 BP; 357 A; 219 C; 407 G; 415 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:      3.63e-165      Length:      1398
Score:          1811.50      Matches:      351
Percent Similarity:      84.0%      Conservative:      48
Best Local Similarity:      73.9%      Mismatches:      57
Query Match:      74.1%      Indels:      19
DB:              6              Gaps:      7

US-10-731-525-8 (1-466) x ADG88184 (1-1398)

Qy      1  MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db      1  ATGGCAATGTCGTAAATGTTCTTCTCTCTCGTCTTCTGGGATCATAAACTCGTTTC 60
Qy      20  -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db      61  GGTGTTTCATTGGAGCCAAAGATTTCGCAATTTGTTTGGTTCGTGAGTTATTGGATCGTGT 120
Qy      36  IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db      121  CATGTTGCTCTGTGTCTCTGAATCTATCTGGGAAGCGATCATCTCTGTTTAAACCTTTA 180
Qy      55  AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db      181  AACGCTGAACCAAGACAAAGAGGATTCAATGATTCTCTTCGCGCAACAATCGTAGCAGAA 240
Qy      75  SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db      241  ATTCAGAGAGGAGTTGAAGTGTGAGATTGAGATTTTGAAGAGCTTGCTAAGAGTTA 300
Qy      92  AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db      301  GAGAAATGCTTCACCTTTGAGATTATGACAAAGCTCTTGAAGAAATACCGGAACGATATC 360
Qy      112  AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db      361  GCCATTGCATTTAGTGGTGCAGAAAGATGTTCTCTTATTAGTAGTACGCTCATTTTGACTGG 420
```


PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136339P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140951P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145227P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.
 Alignment Scores:
 Pred. No.: 4,82e-165 Length: 1712
 Score: 1811.50 Matches: 351
 Percent Similarity: 84.0% Conservatives: 48
 Best Local Similarity: 73.9% Mismatches: 57
 Query Match: 74.1% Indels: 19
 DB: 3 Gaps: 7
 US-10-731-525-8 (1-466) x AAC47286 (1-1712)
 QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
 DB 67 ATGGCAATGCTGTAAATGTTCTTCTTCTCGTCTTCGGGATCATAAACTCTCGTTTC 126
 QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
 DB 127 GGTGTTTCATTTGGAGCCCAAGATTTCGCAAAATGGTTCTGTGAGTTATTGGATCGGTT 186
 QY 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
 DB 187 CATGTTGCTCTGTGCTCTGCAATCATCTCTGGAGGATCATCATCTGTTAAACCTTTA 246
 QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
 DB 247 AACGCTGAACCAAGCAAGGATTCATGATTCCTCTTGGCGCAACAATGGTAGCAGAA 306
 QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
 DB 307 ATTGCAGAGAAGTTGAAGTGGTTGAGATTGAGGATTTTCAAGAGCTTGTCAAGAAGTTA 366
 QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
 DB 367 GAGNATGCTTCACTCTTGAGATTATGGACAAAGCTCTTGAGAAATACGGGACGATATC 426
 QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
 DB 427 GCCATTGCAATTAGTGTGAGAGATGTTGCTCTTATTGAGTAGCGTCTATTGACTGGG 486
 QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
 DB 487 AGGCCATTAGAGTATTAGTTGGATACAGGGAGGTTGAATCCTGAGACGTATCGGTTT 546
 QY 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
 DB 547 TTGATGCGGTGGAGAGCACTATGGATTAGATTAGATTATGATTGTTCTGATTCGTT 606
 QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
 DB 607 GAGGTTCAAGGTTTGGTTAGGACCAAGGATGTTCTCTTTTATGAGGATGTCATCAG 666
 QY 192 GluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAla 211
 DB 667 GAGTGTGCGGTGTTCAAAAGGTGAGACCTTTGAGCGTGCTCTCAAGGGTTTTAAAGGCT 726
 QY 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
 DB 727 TGGATTACTGTCAGAGAAAGATCAATCTCCGGGCAAGGCTCTGAGATTCGGTTGTT 786
 QY 232 GlnValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTrpAsn 251
 DB 787 CAGTTGATCCGGTGTGTTGAAGTTTGGATGGTGGAGTTGTTGTTGGTGAAGTGAAT 846
 QY 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
 DB 847 CCGGTTGCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTGGAGGACTATGGATGTTCCG 906

QY 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
 DB 907 GTTAAACACATTGATCGCGCAGGGTATATATCATGATTGGATGTGAGCCCTTCACGAAAGCG 966
 QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLys 311
 DB 967 GTTTTACCGGTCAGCACGAGAGAGGAGATGTTGGTGGGAAGATGCTAAAGCCCAAG 1026
 QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsn 331
 DB 1027 GAATGTGACTTCACAAAGGGAATGTCAAA---GAAACCTCCGATGATGCTAAAGTGAAC 1083
 QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
 DB 1084 GGGGAATCGAAATCC-----GCTGTTCAGATATCTTTAAGAGT 1122
 QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
 DB 1123 GAGAAATCTTGTGACTTTGAGCAGGCGAGGATTCAGAAATTTGATGAAGTTGGAGAACCGT 1182
 QY 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
 DB 1183 AAAGAGCCTTGGATCGTGTGCTTTATGCTCCGTGGTGCCTTTTGTCAAGCCATCGAA 1242
 QY 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
 DB 1243 GCATCGTATGATGAACCTGGCGGATTAATTTGGCTGGAAGT---GGATTAGGTTGCCAAA 1299
 QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
 DB 1300 TTCAAGCAGATGCTGACCAAGAGGATTTGCTTAAGCAGGAATTGCAGCTCGGTAGCTTC 1359
 QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
 DB 1360 CCTACCATTTCTGTTTTCCTTAAGAACTCATCGAGACCG---ATCAAGTATCCGTCTGAG 1416
 QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
 DB 1417 AAGAGAGATGTTGAGTCTTTGACTTCTGTTCTTGAATCTTGTCGA 1461
 RESULT 9
 AAC37691
 ID AAC37691 standard; DNA; 1714 BP.
 XX
 AC AAC37691,
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18317.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000BP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 25-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135114P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144322P.
PR 20-JUL-1999; 99US-0144484P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0030101439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130851P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135112P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 22-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 4,96e-164 Length: 1702
 Score: 1801.00 Matches: 348
 Percent Similarity: 84.9% Conservative: 51
 Best Local Similarity: 74.0% Mismatches: 55
 Query Match: 73.7% Indels: 16
 DB: 3 Gaps: 6

US-10-731-525-8 (1-466) x AAC38867 (1-1702)

QY 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThrPhe 16
 DB 78 ATGGCACTAGCAATCAACGTTTCTTCATCTCTCTTCTTCGCGATCTCAAGCTCTAGCTTC 137
 QY 17 ProSerSerProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIle 36
 DB 138 CCTTCTTCAGATCTCAAGTAACAAANATCGGATCATAGGTTATTGAATCGTACCAT 197
 QY 37 GlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAla 56
 DB 198 GTCTCTCGGGCTTCTCTGAGTTTGTCCGGAGAGATCTCCCTGAAAGCTCTTAATGTG 257
 QY 57 GluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAla 76
 DB 258 CAATCAATTAACAAGGAATCCATTGTT-----GCTTCTGAGGTTACAGAGAAGCTA 308
 QY 77 SerGluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerPro 96
 DB 309 GATGTGTTGGAAGTTGAGAACTTTGAGAACTTAGCAAGAGATTAGAGAAATGCTTCTCT 368
 QY 97 LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer 116

DB 369 CTTGAGATCATGTATAAAGCTCTTGAGAAGTTTGGGAATGACATTTGCTATTGCTTTAGT 428
 QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
 DB 429 GGAGCTGAAGATGTTGCTCTGATTGAGTATGCTCATTTAACTGGAAGACCTTTATAGGGTA 488
 QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu 156
 DB 489 TTCAGTTTGGATACAGGAGATTGAATCCAGAAACATATAGACTCTTCGATACCGTCGAG 548
 QY 157 LysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeu 176
 DB 549 AAACATTACGGTATTCCGATTGAGTATATGTTTCTCTGATGCTGTTGAGGTTCAAGCTTTA 608
 QY 177 ValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgVal 196
 DB 609 GTTAGGAACAAGGGTTTGTCTCTTCTATGAAGATGGTCACAGGAGTGTGCGGTATT 668
 QY 197 ArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGln 216
 DB 669 AGAAGGTGAGACCTTTGAGGCGTGGTTGAAGGGTTTACGTGCTTGGATCACTGGTCAG 728
 QY 217 ArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAla 236
 DB 729 AGAAAGATCAATCACCCGGGACAGATCGAGATTCTCTGTTGTCAGGTTGATCCAGTG 788
 QY 237 PheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnVal 256
 DB 789 TTTGAAGGATTAGATGGTGAGTTGGTAGTTTGGTGAAGTGAATCCGGTTGCTTAATGTT 848
 QY 257 LysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHis 276
 DB 849 GAAGGGAATGATGTTTGGAAATTTCTTGAGGACCATGGATGTTCTGTCAACACGCTTCAC 908
 QY 277 AlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGln 296
 DB 909 GCTGCGGGTTATGTTTCAATCGGATGTGAGCCATGCACGAGAGCGGTTTTGCGGGTCAG 968
 QY 297 HisGluArgGluGlyVargTrpTrpGluAspAlaLysGlyCysGlyLeuHis 316
 DB 969 CATGAGAGAGAGGGAGATGGTGGGGAAGATGCTAAGGCTAAGAGTGTGCGCTTCAC 1028
 QY 317 LysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSer 336
 DB 1029 AAAGGGAATCAAGGAG-----ATATACCAATGGAAACGCAACCGCT 1070
 QY 337 HisAlaAsnGlyAspAlaThrValProAspIlePheAsnSerProAsnValValAsn 356
 DB 1071 AATGTCAACGGGACAGCC---ACAGTCCGGGATATCTTCAATAGCGAGAAATGTTGTGAAC 1127
 QY 357 LeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeu 376
 DB 1128 TTGACAGCAAGGGAATTGAGAAATTGATGAAGTTAGAGAACCGGAAGAGGCTTGATTC 1187
 QY 377 ValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAsp 396
 DB 1188 GTTGTGCTTTATGACCGTGGTGCCGCTTTGTGTCAGCAATGGAAGCTTCATTTCGATGAA 1247
 QY 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416
 DB 1248 TTGGCGGATAAGCTGGGAGGAAGT---GGTGTGAAGGTTGCTAAGTTTAGAGCTGATGGT 1304
 QY 417 GluGlnLysGlyPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhe 436
 DB 1305 GACCAAGAGGACTTCCAAAAAGGAGTTGACGCTTGGAAAGCTTTCACAACTACTTGTG 1364
 QY 437 PheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArgAspValAsp 456
 DB 1365 TTCCCAAGAACTCTTCAAGACCA---ATCAAGTATCCATCAGAGAGAGAGATGTTGAT 1421
 QY 457 SerLeuMetAlaPheValAsnAlaLeuArg 466

Db 1422 TCITTGACATCTTCTTGAACTCTGTTCCG 1451
RESULT 11
AAC43415
ID AAC43415 standard; DNA; 1377 BP.
XX
AC AAC43415;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 39173.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144894P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147132P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.


```
Db 1111 GTTGTGCTTTATGACCGTGTGCGCGTTTGTTCAGGCAATGGAAGCTTCATTTCGATGAA 1170      :
Qy 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416      :
Db 1171 TTGGCGGATAGCTGGGAGGAAGT--GGTGTGAAGGTTCTTAAGTTTAGAGCTGATGGT 1227      :
Qy 417 GluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhe 436      :
Db 1228 GACCAGAAGGACTTTCACAAAAGGAGTTGCAGCTTGGAAAGCTTTCACAAATACTTGTG 1287      :
Qy 437 PheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArgAspValAsp 456      :
Db 1288 TTCCCAAGAACCTTTCACAGCCA---ATCAAGTATCCATCAGAGAAGAGAGATGTTGAT 1344      :
Qy 457 SerLeuMetAlaPheValAsnAlaLeuArg 466      :
Db 1345 TCATTGACATCTTCTTGAATCTTGTTCGG 1374      :

RESULT 12
AAC32729
ID AAC32729 standard; DNA; 1687 BP.
AC AAC32729;
XX
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 432.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-01231180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 21-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 22-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
```

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 6,36e-160 Length: 1687
Score: 1758.50 Matches: 353
Percent Similarity: 83.2% Conservative: 43
Best Local Similarity: 74.2% Mismatches: 49
Query Match: 71.9% Indels: 31
DB: 3 Gaps: 10

US-10-731-525-8 (1-466) x AAC32729 (1-1687)
QY 1 MetAlaLeuAlaPheThrSerIleSerAlaProThr---SerThrPhe----- 16
DB 98 ATGGCTTTAGCTGTTACTTCTTCAACTGCAATCTCTGATCGACTTCTCTCGTTCT 157
QY 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
DB 158 GGAGCTTCTTCTGAATCTAAAGCTCTTCAATATGTTGATTAGGTTATCTGATCGAAC 217
QY 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55
DB 218 -----CATTGTCCTCAGAGAGCTTACTCTATGAACCCCTTAAC 256
QY 56 AlaGluPro-----ProArgLysAspSerIleValProLeuAlaAlaThrIleVal 73
DB 257 GCTGAGTCATCATACAGACGAATCTTGGGTT--CTTCGTGCTTCAACTCTAATT 313
QY 74 AlaSerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAsp 90
DB 314 GCTCTCTGAAGTTTGAAGAGAAAGAGAGAGTTGAAGACTTTGAGCAACTTGTAAAG 373
QY 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
DB 374 CTTGAAGATGCTTCTCCACTTGAATCATGGATAAAGCTCTTGAGAAATTTGAGAGCAA 433
QY 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluIleAlaLysLeuThr 130
DB 434 ATCGCAATTGCTTTTAGTGAGGCTGAAGATGTTGCATTGATGAATATGCTCGTTAACT 493
QY 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyGln 150
DB 494 GGAAGCCATTTAGGGTTTTTAGTTTAGATACAGGAGAGATTAACCCGGAACGTCACAG 553
QY 151 LeuPheAspAlaValGluLysHisTyGlyIleArgIleGluIleThrMetPheProAspAla 170
DB 554 CTCTTTGACGCGAGTCGAGAGAGCAGTACGGGATTCGAATTGAGTACATGTTTCTTGATGCA 613
QY 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyArgSerGlyHis 190
DB 614 GTTGAGGTTCAAGCTTTAGTGAGAACAGGGTTTGTCTCATCTCTGAAGATGGTCAT 673
QY 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
DB 674 CAAGAGTTGCGCTGTGAGAAAGTTAGACCTTTGCGTGTCTCTTAAGGGTCTTAAA 733
QY 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
DB 734 GCTTGATTACAGGACAGAGAAAGACCAATCTCCGGGTACGAGATCTGAGATCCCTATT 793
QY 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
DB 794 GTTCAGGTTGATCCAGTGTTTGAGGGTTAGATGGCGGTGTGGAAGTCTTGTGAGTGG 853
QY 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
DB 854 AATCCTTTGGCTAAATGTTCAAGGAGCTGATGTGGAACCTTTCTGAGAACTATGGATGT 913
QY 271 ProValAsnSerLeuHisAlaLysGlyTyValSerIleGlyCysGluProCysThrArg 290
DB 914 CCGGTGAATGCATTGTCATGCAAGGATATGTGTCATTCGCGGTGTGAGCCGTGTACTAGG 973
```


QY 409 ValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeu 428
 DB 55387 GTTCCCAAAATTCAGAGCAGATGTCACAGAGAGTTTCTTAAGCAGGAATTCGAGCTC 55446
 QY 429 GlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyr 448
 DB 55447 GGTAGCTTCCTACCAATTCCTGGTTTCCCTTAAGAACTCATCGAGACCG---ATCAAGTAT 55503
 QY 449 ProSerGluLysArgAspValAspSerLeuMetAlaPheValAlaLeuArg 466
 DB 55504 CCGTCTGAGAGAGAGATGTTGAGTCTTTCGACTTCGTTCTTGAATCTTGTCGCA 55557

RESULT 14

AB214500
 ID AB214500 standard; DNA; 1365 BP.

XX AC AB214500;
 XX DT 21-JAN-2003 (first entry)
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2305.
 XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX OS Arabidopsis thaliana.
 XX PN WO200216655-A2.
 XX PD 28-FEB-2002.
 XX PF 24-AUG-2001; 2001WO-US026685.
 XX PR 24-AUG-2000; 2000US-0227866P.
 XX PR 26-JAN-2001; 2001US-0264647P.
 XX PR 22-JUN-2001; 2001US-0300111P.
 XX PA (SCRI) SCRIPPS RES INST.
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Harper JF, Kreps J, Wang X, Zhu T;
 XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 2305; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

XX SQ Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,57e-158	Length:	1365
Score:	1738.00	Matches:	346
Percent Similarity:	82.5%	Conservative:	45
Best Local Similarity:	73.0%	Mismatches:	55
Query Match:	71.1%	Indels:	28
DB:	6	Gaps:	9

US-10-731-525-8 (1-466) x AB214500 (1-1365)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
 DB 1 ATGGCTTTAGCTGTTACTTCTTCTTCAACTGCAATCTCTGGATCGAGTTTCTCTCGTTCT 60
 QY 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
 DB 61 GGAGCTTCTTCTGAATCTAAAGCTTCTCAATATGTTGATAGGTATGATCGAACC 120
 QY 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55
 DB 121 -----CAITGTTCTCAGAGAGCTTACTCTATGAAACCTTTTAAAC 159
 QY 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
 DB 160 GCTGAGTCACATTCACGAAGCAATCTTGGGTTTACTCGTCTTCAACTCTAATTTGCTCCT 219
 QY 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
 DB 220 GAAGTTTGAAGAGAAAGGAGGAGAAAGTTGAAGACTTTTGAGCAACTTGTCTAAAGAGCTTGA 279
 QY 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
 DB 280 GATGCTTCTCCACTTGAATCATGATTAAGCTCTTGAGAGATTCGGAGACCAATCGCA 339
 QY 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
 DB 340 ATTGCTTTTAGTGGAGCTGAAGATGTTGCATTCATTGAATATGCACGCTTTAAGCTGGAAG 399
 QY 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
 DB 400 CCATTTTAGGGTTTTTAGTTAGATACAGGAGAGATTAAACCTCTGAAACGTCAGGCTCTTT 459
 QY 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
 DB 460 GACCGAGTCCAGAGCAGTACGCGATTCGAATTTAGTACATGTTTCTGTAGCATTTGAG 519
 QY 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
 DB 520 GTTCAAGCTTTAGTGAGGAACAAGGGTTGTTCTCTCATTCTATGAAGATGTCATCAAGAG 579
 QY 193 CysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTyr 212
 DB 580 TGTTCGCGTGTGAGGAAGTTAGACCTTTGCGTGTGCTCTTAAGGGTCTTAAAGCTTGG 639
 QY 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
 DB 640 ATTACAGGACAGAGGAAGAACCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCA 699
 QY 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsnPro 252
 DB 700 GTTGATCCAGTGTTCGAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 759
 QY 253 ValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProVal 272
 DB 760 TTGGCTTAATGTTGAGAGAGCTGATGTGTGGAACCTTTCTGAGAACTATGATGTTCCGGTG 819
 QY 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
 DB 820 AATGCATTCACGACACAGAGGTATGTCATCGGGTGTGAGCGGTGTACTAGCCCGGTG 879
 QY 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 312
 DB 880 CTTCCAGGCCAACATGAG 939
 QY 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332
 DB 940 TGTGTGCTACACAAAGGGAACATC-----AAGGAGGAGAGAT----- 975
 QY 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352
 DB 976 -----GGTGTCTGCAGACTCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026

QY 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
 Db 1027 AATGTGGTGCATGTAGCAAGAGGGGTGTGAGAACTTTTGAAGCTAGAGAACCGTA 1086
 QY 373 GluProTriLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGlu 392
 Db 1087 GAGCGTGGTGTGCTGCTTACGCTCCTTGTGTGCCCTTTCGCCAGGGCGATGGAAGCA 1146
 QY 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412
 Db 1147 TCGTACATCAATTTGGCTGAGAACTTTCGGGAAAA---GGAGTTAAAGTGGCGAAATTC 1203
 QY 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
 Db 1204 CGAGCTGACGCTGACGACAGAGGAGTTTGTGAAGCAAGAGCTTCAGTTAGGGAGCTTCCCG 1263
 QY 433 ThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLys 452
 Db 1264 ACATACTTCTCTTCGAAAAAGAGCTCCACGG---GCTATTAAAGTACCCCTTCAGAGCAT 1320
 QY 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
 Db 1321 AGAGATCTGATTCATCACTGCTGCTTGTGTAATCTTCTTCGG 1362
 RESULT 15
 ADG88214
 ID ADG88214 standard; cDNA; 1365 BP.
 XX
 AC ADG88214;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE A. thaliana RPP4-upregulated pathogen infection-related gene #656.
 XX
 KW Pathogen infection-related gene; plant; Peronospora parasitica;
 KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
 KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200222675-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 14-SEP-2001; 2001WO-US028506.
 XX
 PR 15-SEP-2000; 2000US-0232778P.
 PR 22-JUN-2001; 2001US-0300183P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA (UYN-) UNIV NORTH CAROLINA.
 PA (GLAZ/) GLAZEBROOK J.
 PA (WANG/) WANG X.
 PA (DANG/) DANG J L.
 PA (EULG/) EULGEM T.
 PA (ZHUT/) ZHU T.
 XX
 PI Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
 XX
 DR WPI; 2002-292409/33.
 XX
 PT Novel isolated polynucleotide, useful for conveying pathogen resistance
 PT to plants, and for identifying plants infected with a pathogen.
 XX
 PS Claim 3; SEQ ID NO 656; 605pp; English.
 XX
 CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
 CC ADG87557) whose expression is altered in response to pathogen infection,
 CC and to homologues of these genes from other plants or fungi, especially
 CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
 CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
 CC expression of genes of the invention was upregulated or downregulated in
 CC Arabidopsis plants infected with the oomycete Peronospora parasitica,

CC indicating that they play a role in defence mechanisms. The genes of the
 CC invention are regulated by RPP7 or RPP8 which act via unconventional
 CC signalling cascades, or by the RPP4-dependent pathway. The invention also
 CC relates to polypeptides encoded by the pathogen infection-related genes;
 CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
 CC ; expression cassettes, host cells and pathogen-resistant transgenic
 CC plants and their progeny comprising a polynucleotide of the invention;
 CC and a method of identifying a plant cell infected with a pathogen. The
 CC polynucleotide sequences and methods of the invention are useful for
 CC identifying plants infected with a pathogen, and for conferring
 CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
 CC nematodes and insects (e.g., aphids). The present sequence represents an
 CC Arabidopsis thaliana gene whose expression is altered in response to
 CC Peronospora parasitica infection. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences.

XX
 SQ Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.57e-158 Length: 1365
 Score: 1738.00 Matches: 346
 Percent Similarity: 82.5% Conservative: 45
 Best Local Similarity: 73.0% Mismatches: 55
 Query Match: 71.1% Indels: 28
 DB: 6 Gaps: 9

US-10-731-525-8 (1-466) x ADG88214 (1-1365)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
 Db 1 ATGCTTTAGCTGTACTTCTTCAACTGCAATCTCTGATCGAGTCTTCTCGTTCT 60
 QY 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
 Db 61 GGAGCTTCTTCTGAATCTAAAGCTTTCAAATATGTTTCGATTAGTTATCTGATGAAC 120
 QY 36 IleGlyGlyAlaValAsnPheAsnLeuSerGluArgArgSerLeuValLysProValAsn 55
 Db 121 -----CATTTGCTCAGAGACGTTTGAAGACTTTTGAGCAACTTGTCTAAAGCTTGA 159
 QY 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaIleThrIleValAla 74
 Db 160 GCTGAGTCACATTCACGAGCGAATCTTGGTTACTCGTCTCACTCACTTAATGCTCT 219
 QY 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
 Db 220 GAAGTTGAAGAGAAAGGAGGAGAGTGAAGACTTTTGAGCAACTTGTCTAAAGCTTGA 279
 QY 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
 Db 280 GATGCTTCTCACCTTGAAATCATGATAAAGCTCTTGAGAGATTCCGAGAGCAAAATCGCA 339
 QY 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
 Db 340 ATTGCTTTTAGTGAGCTGAAGATGTTGCAATTGATTAATGACGCTTTAACTGGAAG 399
 QY 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
 Db 400 CCATTATGGGTTTTTAGTTAGTACAGGGAGATTAAACCCCTGAAACGTACAGGCTCTTT 459
 QY 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
 Db 460 GACGAGTCGAGAGACGAGTACGGGATTTCGAATTGAGTACATGTTTCTGATGCAGTTGAG 519
 QY 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
 Db 520 GTTCAAGCTTTAGTGAGGAGCAACAGGGTTTGTCTCATCTTATGAAGATGGTCATCAAGAG 579
 QY 193 CysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrp 212
 Db 580 TGTTCGCGTGAGGAAAGTTAGACCTTTGCGTCTGCTCTTAAAGGGTCTTTAAAGCTGG 639

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 20, 2006, 22:17:46 ; Search time 5571 Seconds
(without alignments)
3913.621 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPTSTPSSSE.....KYPSEKRDVDSLMFAFNALR 466

Scoring table:

	BLOSUM62	Xgapext	0.5
Xgapop 10.0	0.5		
Xgapop 10.0	0.5		
Xgapop 6.0	7.0		
Delop 6.0	Delext 7.0		

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB.spool/US10731525/runat.17022006.164858.123/app_query.fasta.1
-DB-EST -QPMT-fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05h
-USRP=US10731525 @CGN 1.1 5315 @runat.17022006.164858.123 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*	1:	gb_est1.*
2:	gb_est2.*	
3:	gb_est3.*	
4:	gb_est4.*	
5:	gb_est5.*	
6:	gb_est6.*	
7:	gb_est7.*	
8:	gb_est8.*	
9:	gb_ges1.*	
10:	gb_ges2.*	
11:	gb_ges3.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1702	69.6	1579	4	CNSOACVY BX815501 Arabidops
2	1689	69.1	1535	4	CNSOADIU BX813913 Arabidops
3	1500	61.3	1299	10	CL972938 OsIFCC023
4	1346	55.1	959	7	CK264757 EST710835
5	1279	52.3	830	6	CB339330
6	1256	51.4	924	7	CV290140 aof01-6ms
7	1241.5	50.8	1308	4	AY105476 Zea mays

8	1226.5	50.2	836	9	BH652825
9	1204	49.2	910	6	CD573900
10	1187.5	48.6	831	10	CW799574
c 11	1172	47.9	764	6	CF119175
c 12	1172	47.9	764	6	CF119189
c 13	1166.5	47.7	968	7	CK263241
c 14	1163	47.6	704	6	CB340603
15	1156	47.3	856	7	CK287268
16	1149	47.0	859	8	DN982440
17	1148.5	47.0	780	3	BM111117
18	1135	46.4	858	2	BF253965
19	1130	46.2	753	5	BQ915480
c 20	1120	45.8	756	10	DU052684
c 21	1114	45.6	741	5	BQ969575
c 22	1111	45.4	755	5	BU025975
c 23	1100	45.0	875	6	CB292552
c 24	1100	45.0	987	8	DR932581
c 25	1093	44.7	829	2	BF623476
c 26	1084	44.3	729	6	CA808814
c 27	1083	44.3	864	10	CG265597
c 28	1083	44.3	963	7	CK266388
c 29	1082	44.3	612	2	BE058404
c 30	1080.5	44.2	753	3	BQ118692
c 31	1077.5	44.1	719	5	BQ994338
c 32	1077.5	44.1	744	7	CV031415
c 33	1076	44.0	728	7	CV098569
c 34	1075.5	44.0	778	8	DN149426
c 35	1074.5	43.9	653	2	BG124326
c 36	1074	43.9	716	7	CO492657
c 37	1073.5	43.9	689	2	BG597316
c 38	1067.5	43.7	775	8	DN149322
c 39	1059.5	43.3	957	9	CC611389
c 40	1057.5	43.3	683	6	CA924320
c 41	1055.5	43.2	676	6	CA932058
c 42	1051	43.0	648	6	CD716959
c 43	1050.5	43.0	691	8	DR399755
c 44	1048	42.9	831	8	CX053400
c 45	1046.5	42.8	671	7	CO105768

ALIGNMENTS

RESULT 1	CNSOACVY	1579 bp	mrna	linear	HTC 06-FEB-2004
CNSOACVY	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
LOCUS	GSRLUS68B02 of Adult vegetative tissue of strain col-0 of				
DEFINITION	Arabidopsis thaliana (thale cress).				
ACCESSION	BX815501				
VERSION	BX815501.1				
KEYWORDS	HTC; GSRL cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1579)				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1579)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.				

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

```

1. .1579
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL682B02"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
complement(1..1579)
/gene="Atlg62180"

```

gene

ORIGIN

```

Alignment Scores:
Pred. No.: 2.33e-167 Length: 1579
Score: 1702.00 Matches: 339
Percent Similarity: 81.6% Conservative: 48
Best Local Similarity: 71.5% Mismatches: 59
Query Match: 69.6% Indels: 28
DB: 4 Gaps: 9

```

US-10-731-525-8 (1-466) x CNSOACVY (1-1579)

```

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
Db 90 ATGGCTTTAGCTGTTACTTCTTCTCAACTGCAATCTCTGGATCGAGTTCTCTCGTCT 149
Qy 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 150 GGAGCTTCTCTGTAATCTAAAGCTCTTCAGATATGTTTCGATTAGTTATCTGATCGAAC 209
Qy 36 IleGlyGlyAlaValAsnLeuSerGlnArgSerLeuValLysProValAsn 55
Db 210 -----CATTTGCTCAGAGAGCTTACTCTGCTTCAACTCTAATTTGCTCCT 248
Qy 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaIleAlaThrIleValAla 74
Db 249 GCTGAGTCACATTCCAGAGCGAATCTTGGGTACTCTGCTTCAACTCTAATTTGCTCCT 308
Qy 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
Db 309 GAACTTGAAGAGAAAGGAGGAGAGATTGAAGACTTTGAGCAACTTGTCTAAGGAGCTTGAA 368
Qy 93 AsnAlaSerProGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
Db 369 GATGCTTCTCCACTTGAGATCATGATAAGGCTCTTGAGAGATTCCGAGACCAATCGCA 428
Qy 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
Db 429 ATTGCTTTAGTGGAGCTGAAGATGTTGCGTTGATTGAGTGTGACGCTTAACTGGAAAG 488
Qy 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
Db 489 CCAITTTGGGGTTTGGGTTTAGATACAGGGAGATTAAACCCCTGAGACATCAGGCTCTTT 548
Qy 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
Db 549 GACGAGTTCGAGAGCAGTACGGGATTCGAATTGATGATCATGTTTCTGATGCGAGTTGAG 608
Qy 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
Db 609 GTTCAAGCTTTAGTAGGAGAAACAGGGTTTGTCTCATTTCTATGAAGATGGTCAACAGAG 668

```

```

Qy 193 CysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTyr 212
Db 669 TGTTCGGTGTGAGAAAGTTAGACCTTTCGTCGTGCTCTTGAGGGTCTTAAAGCTTGG 728
Qy 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
Db 729 ATTACAGGACAGAGAAAGACCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCA 788
Qy 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsnPro 252
Db 789 GTTGTATCCAGTGTTTGAAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 848
Qy 253 ValAlaAsnValLysGlyHisAspIleThrAsnPheLeuArgThrMetAsnValProVal 272
Db 849 TTGGCTAATGTTGAGGAGCTGATGTGGAATCTTCTGAGAACTACGGATGTTCCGGTG 908
Qy 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
Db 909 AATGCATTGCACGCACACAAAGGTTATGTCAATCGGTTGTGAGCGCTGTACTAGGCCGTG 968
Qy 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLysGlu 312
Db 969 CTTCCAGGCCAACATGAGAGAGAGGAGTGTGTGGTGGGAAGATGCTAAAGCTAAAGAA 1028
Qy 313 CysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsnGly 332
Db 1029 TGTGTTCTACAAAGGGAACATC-----AAGGAGGAGAT----- 1064
Qy 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsnSerPro 352
Db 1065 -----GGTGTCTGCAGACTCAAAGCCTGCTGTCTGCAAGAGATATTGGAAGCAAC 1115
Qy 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
Db 1116 AATGTGTTTCATTCGACAAAGAGGGGTTGAGAATCTTTTGAAGCTAGAGAACCGTAAA 1175
Qy 373 GluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGlu 392
Db 1176 GAGCGTGGTGTGCTGCTTACCTTTCGTCCTTTCGTCGCTTTCGAGCGATGGAAGCA 1235
Qy 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412
Db 1236 TCGTACATCAATGTTGCTGAGAACTTTCGCGGAAAA---GGAGTTAAAGTCGCGAAATTC 1292
Qy 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
Db 1293 CGAGCTACGCGTGAAGAGAGGATTTGCTAAGCAAGAGCTTTCAGTATGAGGAGCTTCCCG 1352
Qy 433 ThrIleLeuPhePheProLysHisSerArgProThrIleLysTyrProSerGluLys 452
Db 1353 ACATATCTTCTATTTCCGAAAAGAGCTCCACGG---GCTATTAGTACCTTTCAGAGCAT 1409
Qy 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1410 AGAGATGTTGATTCATCATGTCGTGTTGTGAATCTTCTTCGG 1451

```

RESULT 2

CNSOADUO

LOCUS

DEFINITION

thaliana (thale cress).

Accession

Version

Keywords

Source

Organism

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1535)

Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

bacterial pathogen Xylella fastidiosa

JOURNAL COMMENT

Unpublished (2003)

Contact: Douglas Cook, PhD

CAES Genome Facility

UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA

Tel: 530 754 6561

Fax: 530 754 6617

Email: drcook@ucdavis.edu

Seq primer: ACGGTACCGACATATGCC.

Location/Qualifiers

FEATURES

source

1..830

/organism="Vitis vinifera"

/mol_type="mRNA"

/cultivar="Cabernet Sauvignon"

/db_xref="taxon:29760"

/clones="CA23EI03IVP B07"

/sex="Hermaphrodite"

/dev stage="Mid-season leaf material, collected July 25, 2001."

/lab_host="DH5alpha"

/clone_lib="Cabernet Sauvignon Leaf - CA23EI"

/note="Organ: Leaf; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CA23EI is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were verified to be infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:

5'-AAGCAGTGATGATCAACGAGTGGCCATACGCCGGG-3' and

5'-ATTCTAGAGCCGAGCGCGCATG-dt(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Alignment Scores:

Pred. No.: 2.43e-123 Length: 830
Score: 1279.00 Matches: 237
Percent Similarity: 92.1% Conservative: 21
Best Local Similarity: 84.6% Mismatches: 16
Query Match: 52.3% Indels: 6
DB: 6 Gaps: 3

US-10-731-525-8 (1-466) x CB339330 (1-830)

Qy 82 GluAspPheGluGlnIleAlaSerAspLeuAspPheAlaSerProLeuGluIleMetAsp 101
Db 9 GAGGACTATGACCAATTGGCTATGACCTTGCATATGCTCTCTCTTGAATATGGAT 68
Qy 102 ArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyValAlaGluAspVal 121
Db 69 AAGGCCCTTGCAAAATTGGCAATGACATTGCCATTGCTTTCAGTGGTGTGAAGATATT 128
Qy 122 AlaLeuIleGluTyrAlaLysLeuThrGlyArgPropheArgValPheSerLeuAspThr 141
Db 129 GCTTTGATTGAATGACCGCTTAACCTGCTCGCTTTAGGGTATTGAGCTGGACACT 188
Qy 142 GlyArgLeuAnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyTle 161
Db 189 GGAAGGTTGAACCCGGAACATATCAATCTTTGACACGGTGTGAACACATATGGCATC 248
Qy 162 ArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGly 181
Db 249 CGCATTGATACATGTTTCCAGATGCTGTTGAGTTTCAGGGCTTGGTGAGGAGCAGGGA 308
Qy 182 LeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArgPro 201
Db 309 CTGTTCTCTCTACAGGATGGGCACAGGAGTGTCTCCGCTGTGAGGAAGGTGAGACCC 368

Qy 202 LeuArgArgAlaLeuLysGlyLeuArgAlaTyrPheThrGlyGlnArgLysAspGlnSer 221
Db 369 TTGAGAGAGGCACTGAAAGGTTTGGCTGCTGATCACAGGGCAAGAGATCACTCT 428
Qy 222 ProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAsp 241
Db 429 CCTGGTACCAGGGCTGAAGTTCCTGTGTCCAGGTGGATCCAGCTTCGAGGGGTTGGAT 488
Qy 242 GlyGlyIleGlySerLeuValLysTyrAsnProValAlaAsnValLysGlyHisAspIle 261
Db 489 GGTGGGGTTGGCAGCCTGGTGAATGGAACCCAGTGGCAATGTGCAGGGCATGGATC 548
Qy 262 TrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrVal 281
Db 549 TGGAAATTCCTCGTGTCTATGAATGTGCTGTGAATTCATTCGAAAGGCTATATT 608
Qy 282 SerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGly 301
Db 609 TCAATTGGGTGTGAGCCATGCCTAGGCTGTACTACCTGGCCAGCATGAGAGAGAAGA 668
Qy 302 ArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLys 321
Db 669 AGATGGTGGTGGGAAGATGCCAGGCCAAGAGTGTGGCTTTCATAAAGGAACCTCAAG 728
Qy 322 GlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAsp 341
Db 729 CAGGAA---CATGGAAACAAATGGAATGGG-----CATGCAATGGAACC 773
Qy 342 AlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGly 361
Db 774 GCC---ACTGTTCTGATCTCTTCGACACCCAGCCAGNNNTGTTTACCTTAAACCGAGCTGGG 830

RESULT 6
CV290140/c
LOCUS
DEFINITION
aof01-6ms4-d06 Aof01 Asparagus officinalis cDNA clone
aof01-6ms4-d06 5', mRNA sequence.
CV290140
CV290140.1 GI:52575140
EST.
VERSION
KEYWORDS
SOURCE
ORGANISM
Asparagus officinalis (garden asparagus)
Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.
REFERENCE
1 (bases 1 to 924)
Gepamphilis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,
Oppenheimer, D., Frohlich, M., Doyle, J., Tanksley, S., Webb, M.,
Leebens-Mack, J., Landherr, L., Ilut, D. and Wall, K.
Generation of ESTs from early male inflorescences of Asparagus
officinalis
Unpublished (2004)
Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pgn.cornell.edu>)
Plate: aof01-6ms4 row: d column: 06
Seq primer: M13F.
Location/Qualifiers
1..924
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="PGN:aof01-6ms4-d06"
/db_xref="taxon:4686"
/clones="aof01-6ms4-d06"
/tissue_type="male inflorescences"

FEATURES
source
1..924

Qy		420	uPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHi	440
			:::	
Db		160	GTTTGTCTACGAAGAATTGCAGCTGGAGTTTTCCGACCATACTGTTGTTACTAAGAA	101
			:::	
Qy		440	sSerSerArgProThrIleLysTyrrProSerGluLysArgAspValaspSerLeuMetal	460
			:	
Db		100	TGGCTCGAGACCG--ATAAAGTATCCATCAGAGAAGAGGGATGTTGCATCTCTGCTTCG	44
			:::	
Qy		460	aPheValAsnAlaLeuArg	465
			:::	
Db		43	ATTTCATCAATGCACCTTAGA	25

161	lleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerIlys	180
919	ATCCACATCGAGTACATGTTTCCAGATCGCGCGAAGTCCATCCCTCGTAAAGCAAA	860
181	GlyLeuPheSerPheTyrGluAspGlyHis-GlnGluCysCysArgValArgIysValAr	200
859	GGCCCTTTTTCCTTCTACGAGGAGCGGCCACCCAGAAGATGCTGCGAGTTTGC	800
200	gProLeuArgArgAlaLeuIysGlyLeuArgAlaTrpIleThrGlyGlnArgIysAspG	220
799	CCCTCTTCGAAGAGCTCTCAAGAGCCTCAGAGCTTGATCACCGETTCACGCCAAGATCA	740
220	nSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMe	240
739	GTCTCCTGGAAACCCGAGCCCAAAATCCCGTAGTCCAAAGTCGACCCCTCGTTTGAAGGGCT	680
240	tAspGlyGlyIleGlySerLeuValIysTrpAsnProValAlaAsnValIysGlyHis	260
679	TCACCGTGGGGCTGGCAGCTTGATTAAATGAACCCGGTTGCAAAATGTCAATGGGAAGA	620
260	pIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaIysGlyTV	280
619	TATTGGAGTTTTTTGAGGAGATGGAGGTTCCGGTGAACCTCGTGCATCTCTCAGGSETTA	560
280	rValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgG	300
559	CGTGTGATTCGATGCGAGCGCGTGCACGAGCGCTGTGCTCCAGGGGCGAGCAGCAGAGAGA	500
300	uGlyVArgTrpTrpTrpGluAspAlaIysAlaIysGluCysGlyLeuHisIysGlyAsnVa	320
499	AGGGAGATGGTGGTGGGAGGATGCAAAAGCCAAAGAGTGGCGGCTTCAAAAGGGCAACAT	440
320	lIysGlnGlnIysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnG	340
439	CTCTCAG-----ACTCAACAGTGAACGGAGATGCTGAGTCTCTTCATTCACTGATGCG	386
340	yAspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgTh	360
385	G-----ACTCTGATTGTTTCCAGACTCTCAGCGCATTTGTGAATCTGTCAAGGCC	338
360	rGlyIleGluAsnLeuAlaIysLeuGluAspArgIysGluProTrpLeuValValLeuTy	380
337	GGGGAATTGAGAAATCTGATGAAGTTTGGAGAGAGAGCAGACCTTGGGTGTTGTGCTTTTA	278
380	rAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspIy	400
277	CGCGCCTTGGTGCAGATTTTTCGCCAGCGGATGGAGAGTTCTTATGTTGAGCTAGCTGAGA	218
400	sLeuAlaGlySerThrGlyMetIysValGlyIysPheArgAlaAspGlyGlyGlnIysG	420
217	GTATTACGGGTTTCG---GATGTTAGGTTTGCAAAGTTTCAGAGCTGATGGGAGTAAAGGC	161

Query Match:	50.8%	Indels:	15
DB:	4	Gaps:	4
US-10-731-525-8 (1-466) x AY105476 (1-1308)			
Qy	172	GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln	191
Db	10	GAGGTGCAGGAGCTGGTGGCCACCAAGGGCCTCTCTCTCTACGAGGACGGCCACGAG	69
Qy	192	GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla	211
Db	70	GAGTGTCTCGGGTGGCGCAAGGTGGGCCCTTGGCAGGGCGCTCAAGGGGCTTAGGGCA	129
Qy	212	TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal	231
Db	130	TGGATCACCGGCCAGAGAAAGACAGTCCCGCGGACAGGCGCAGCATCCCCATTTGTC	189
Qy	232	GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn	251
Db	190	CAGGTTGATCTCTTCGAGGCTGGATGGCGGGCGCGTAGCTTGGTCAAGTGAAC	249
Qy	252	ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro	271
Db	250	CCCGTGGCCACAGTCGACGCAAGGACATCTGGACTTTCTCGGACCATGGACGTACCT	309
Qy	272	ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro	291
Db	310	GTCAACACCTGCTCCTCAGGGCTACGTGTCCATCGGGTCCGAGCCGTGCACAGGCC	369
Qy	292	ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLys	311
Db	370	GTCTCTCGGGGCGAGCAGCAGCGTGAAGGCGGTGTGTGGAGGACGCCCAAGGCCAAG	429
Qy	312	GluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGlu-----GluAsp	327
Db	430	GAGTGGCGGCTCCACAAGGGCAACATTGACAAGGACGCCCGGCGGGCGGCCCGAGTCC	489
Qy	328	ValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAsp	347
Db	490	GCCAACGCCAACGGC-----TCGGCGGGCGGCCCGGAC	522
Qy	348	IlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLys	367
Db	523	ATCTTCGAGAGCCCGCGGTGTGTCTCCCTCACCGCAGCGGATCGAGAACCTGTGCGC	582
Qy	368	LeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCys	387
Db	583	CTGGAGAACCGCGCGAGCCGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	642
Qy	388	GlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMet	407
Db	643	CAGGCCATGAGGACCTCTCTACGTGGAGCTGGCGAGAGCTGGCGGGTCC---GGGTG	699
Qy	408	LysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGln	427
Db	700	AAGGTGGCCAAAGTTCGCGCGGAGCGCGAGCAGAACCGCTTCGCGCAGGCCGAGTGCAG	759
Qy	428	LeuGlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLys	447
Db	760	CTGCAGAGCTTTCCACCGCTGTCTGTTCGGCGGGCGCCAGCCCGCAGGCC---ATCAAG	816
Qy	448	TyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg	466
Db	817	TACCCGTGGAGAGAGGGACGTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	873
RESULT 8			
BH652825			
LOCUS			
DEFINITION			
BOHYM08TR_BO_2_3_KB Brassica oleracea genomic clone BOHYM08,			
genomic survey sequence.			
ACCESSION			
BH652825			
VERSION			
BH652825.1 GI:18710905			
KEYWORDS			
GSS.			


```

Db      557 GCATTAGTGGAGCAGGGGCTCTTCTCTTTATGAGGATGGGCACAGGAGTGCTGC 616
Qy      195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
Db      617 CGTGTGAGGAAGTGAAGCCCTTCGGAGGGCCCTCAAGGGATTCGCTGCTGATCACT 676
Qy      215 GlyClnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
Db      677 GGCCAAAGGAAGATCATCTCCAGGTACTAGGTCTGAAATCCCTGTAGTCCAGGTGGAT 736
Qy      235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254
Db      737 CCTGTTTTGAGGGCTGGAAGTGGGGTGGCAGCTCGTGAAGTGAATCCAGTGGCT 796
Qy      255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db      797 AATCGAAGGGAATGACATTTGGAACTTCTTCGTACTATGGGTGAGCCCATCAATTC 856
Qy      275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
Db      857 TTGCACTCAAAGGGTACATCTCGATTGGGTGTGAGCCATGCACAAGGCCCGTC 910

```

RESULT 10

CW799574 831 bp DNA linear GSS 23-NOV-2004
 WiscDblx339B11 Arabidopsis thaliana T-DNA insertion flanking
 sequences Arabidopsis thaliana genomic, genomic survey sequence.

ACCESSION

VERSION CW799574

KEYWORDS

SOURCE GSS.

ORGANISM

Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 831)
 Woody, S.T., Monson, S.S., Jester, P.J., Austin-Phillips, S.,
 Anasino, R.M., Sussman, M.R. and Krysan, P.J.

A New Community Resource for Knocking-Out Small Genes and
 Tandemly-Duplicated Gene Families and for Mosaic Analysis in
 Arabidopsis

Unpublished (2004)

Contact: Woody ST

Biotechnology Center

University of Wisconsin-Madison

425 Henry Mall, Madison, WI 53706, USA

Tel: (608) 262-4640

Email: swoody@facstaff.wisc.edu

Class: TAIL-PCR.

Location/Qualifiers

1. 831

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/cultivar="Col-O ecotype"

/db_xref="taxon:3702"

/tissue_type="seeds produced by primary (Baeta-resistant)

transformants"

/clone_lib="Arabidopsis thaliana T-DNA insertion flanking

sequences"

/note="Vector: pBS-Lox; Sequence generated in the course

of an Arabidopsis T-DNA tagging program. TAIL-PCR was used

to generate sequencing templates that represent A.t.

genomic DNA flanking the left border of the pBS-Lox T-DNA

insert. PCR products were sequenced directly by using the

p745 primer 5' AACGTCGCAATGTGTATTAAAGTTGTC 3'

ORIGIN

Alignment Scores: 9.93e-114 Length: 831
 Pred. No.: 1187.50 Matches: 225
 Score: 86.1% Conservative: 23
 Percent Similarity: 86.1% Mismatches: 28
 Best Local Similarity: 78.1%

```

Query Match: 48.6% Indels: 14
DB: 10 Gaps: 3

US-10-731-525-8 (1-466) x CW799574 (1-831)

Qy      148 ThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhe 167
Db      5 ACNTACAGGCTCTTTCAGCCANTCGAGAAGCAGTACGGGATTGGAATTGATGATCATGTTT 64
Qy      168 ProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGlu 187
Db      65 CCTGATGCGATTGAGGTTCAAGCTTTAGTCAGGAACAAGGGTTGTCTCATCTTATGAA 124
Qy      188 AspGlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLys 207
Db      125 GATGGTCATCAAGAGTGTGCCGTGTAGGANAAGTTAGACCTTGGCTTGGCTGCTCTTAAG 184
Qy      208 GlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGlu 227
Db      185 GGTCTTAAAGCTTGGATTACAGGACAGAGGAAAGACCAATCTCCGGGTACGAGATCTGAG 244
Qy      228 IleProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeu 247
Db      245 ATCCCTATTGTTCCAGGTTGATCCAGTGTTTGAAGGGTTAGATGGCGGTGTGGAAGTCTT 304
Qy      248 ValLysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThr 267
Db      305 GTGAAGTGGAACTCTTGGCTATGTTGAAGGAGCTGATGTGGAACCTTCTGAGAACT 364
Qy      268 MetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluPro 287
Db      365 ATGGATGTTCCGGTGAATGATTCAGCCACCAAGGGGTATGTGCAATCCGGGTGTGAGCCG 424
Qy      288 CysThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTropGluAsp 307
Db      425 TGTACTAGCGCGGTGCTTCAGGCCCAACATGAGAGAGAAGAGGTGGTGGTGGGAAGAT 484
Qy      308 AlaIleValLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAsp 327
Db      485 GCTAAAGCTAAAGATGTGCTTACACAAGGGAACATC-----AAGGAGGAAGAT 535
Qy      328 ValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAsp 347
Db      536 -----GGTGTGCAGACTCAAAGCCTGCTGCTGCAAGAG 571
Qy      348 IlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLys 367
Db      572 ATATTTGAAAGCAACAATGTTGGTTGCATTGAGCAAGAGGGGGTGTAGAATCTTTTGAAG 631
Qy      368 LeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCys 387
Db      632 CTAGAGAACCGTAAAGAGCGGTGGTGGTGTACTTTTACGCTCTCTTGGTCCCTTTCTGC 691
Qy      388 GlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMet 407
Db      692 CAGCGCATGGAAGCATCGTACATCGAATTGGCTGAAAAAAGCTTGGCGGAAAA--GGAGTT 748
Qy      408 LysValGlyLysPheArgAlaAspGlyClnLysGluPheAlaLysSerGluLeuGln 427
Db      749 AAAGTGGCGAAATTCNAGCTGACGGGTGAGCANAAGGA-NTTGTCTAACCANAN-CTTCAG 806
Qy      428 LeuGlySerPheProThrIleLeu 435
Db      807 TTAGGGAGCTTCCGACNATCTCTN 830

```

RESULT 11

CF119175/c
 LOCUS CF119175 764 bp mRNA linear EST 23-JUL-2003
 DEFINITION MTU10CS.P14.D12 Aspen stem cDNA Library Populus tremuloides cDNA,
 mRNA sequence.
 ACCSSION CF119175
 VERSION CF119175.1 GI:33183854
 KEYWORDS EST.

```

SOURCE
ORGANISM
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
1 (bases 1 to 764)
AUTHORS
Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
Tsai,C-J.
TITLE
Expressed sequence tags from Aspen
JOURNAL
Unpublished (2003)
COMMENT
Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES
source
Location/Qualifiers
1..764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/note="Organ: stem"

ORIGIN
Alignment Scores:
Pred. No.: 3,73e-112 Length: 764
Score: 1172.00 Matches: 213
Percent Similarity: 90.6% Conservative: 18
Best Local Similarity: 83.5% Mismatches: 22
Query Match: 47.9% Indels: 2
DB: 6 Gaps: 2

US-10-731-525-8 (1-466) x CF119175 (1-764)

QY 190 HisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeu 209
Db CATCAGAGTGCTGCCGTGTAGAGGTGAGACCTTGAGCGGCTCTGAGGGGCTT 705
QY 210 ArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIlePro 229
Db CGGGCTTGGATCATCTGCCAAAGGAAGGATCAATCTCCAGGACGAGGTCTGAATTTCCA 645
QY 230 ValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLys 249
Db GTTGTTCAGGTGGACCCGGTTTTTGAGGGATTGGATTGGTGGGCTGGAAGCCTGTACAAG 585
QY 250 TrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
Db TGGATCAATGCGCAATGTTGAGGACCAAGATGATGGAAGTCTCTCGNACCATGGAT 525
QY 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
Db GTCCCTGTGAACCTCAATTCGATTCATAAGGGATATCTGATTCGAGGCTTGCACAC 465
QY 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
Db AGGCCAGTTCTTCCGGCGCCCAACATGAGAGAAAGGAGTGGTGGTGGGAGGATGCCACA 405
QY 310 AlaLysGluCysGlyLeuHisLysGlyAsnValLysGlyGlnLysGluGluAspValAsn 329
Db GCTAAGGAATGTGGTCTTCATAAAGGAATCTGAAACAGGGTATCGACGCCCAACTTAT 345
QY 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
Db CGCAATGGAAACCGGGCTGCCCATGCCAACCGG---GCTGCCACTGTGTGTGATATTTTC 288
QY 350 AsnSerProAsnValValLeuLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu 369
Db AACTCCGAGAATTTGGTCAATTTTGGACGGCCCGGAATCGAGAACTTATTTAAATATTAGAG 228

```

```

QY 370 AspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAla 389
Db AACCGAAAGGAACCATGGCTTGTGGTCTCTACGCCCATGGTCCCAATTTTCCCAAGGT 168
QY 390 MetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysVal 409
Db ATGGAAGCATCTTATGTTGAATTTGGCCGATAGTTAGCAGGGAGT---GGAATAAAGGTG 111
QY 410 GlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGly 429
Db GGAAATTCAGGCACATGGCATCAGAGGAATTTCTCAAAACAAGATTTACAGCTCGGA 51
QY 430 SerPheProThrIleLeuPhePheProLysHisSerSerArgPro 444
Db AGTTTTTCCAACAATCTTTTCTTCTTCCCTAAACACTCATCTCGCCCA 6
RESULT 12
CF119189/c
LOCUS
DEFINITION
764 bp mRNA linear EST 23-JUL-2003
MTU10CS.P14.F03 Aspen stem cDNA Library Populus tremuloides cDNA,
mRNA sequence.
ACCESSION
CF119189
VERSION
CF119189.1 GI:33183882
KEYWORDS
EST.
SOURCE
Populus tremuloides (quaking aspen)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
1 (bases 1 to 764)
AUTHORS
Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
Tsai,C-J.
TITLE
Expressed sequence tags from Aspen
JOURNAL
Unpublished (2003)
COMMENT
Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES
source
Location/Qualifiers
1..764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/note="Organ: stem"

ORIGIN
Alignment Scores:
Pred. No.: 3,73e-112 Length: 764
Score: 1172.00 Matches: 213
Percent Similarity: 90.6% Conservative: 18
Best Local Similarity: 83.5% Mismatches: 22
Query Match: 47.9% Indels: 2
DB: 6 Gaps: 2

US-10-731-525-8 (1-466) x CF119189 (1-764)

QY 190 HisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeu 209
Db CATCAGAGTGCTGCCGTGTAGAGGTGAGACCTTGAGCGGCTCTGAGGGGCTT 705
QY 210 ArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIlePro 229
Db CGGGCTTGGATCATCTGCCAAAGGAAGGATCAATCTCCAGGACGAGGTCTGAATTTCCA 645
QY 230 ValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLys 249
Db GTTGTTCAGGTGGACCCGGTTTTTGAGGGATTGGATTGGTGGGCTGGAAGCCTGTACAAG 585
QY 250 TrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
Db TGGATCAATGCGCAATGTTGAGGACCAAGATGATGGAAGTCTCTCGNACCATGGAT 525
QY 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
Db GTCCCTGTGAACCTCAATTCGATTCATAAAGGGATATCTGATTCGAGGCTTGCACAC 465
QY 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
Db AGGCCAGTTCTTCCGGCGCCCAACATGAGAGAAAGGAGTGGTGGTGGGAGGATGCCACA 405
QY 310 AlaLysGluCysGlyLeuHisLysGlyAsnValLysGlyGlnLysGluGluAspValAsn 329
Db GCTAAGGAATGTGGTCTTCATAAAGGAATCTGAAACAGGGTATCGACGCCCAACTTAT 345
QY 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
Db CGCAATGGAAACCGGGCTGCCCATGCCAACCGG---GCTGCCACTGTGTGTGATATTTTC 288
QY 350 AsnSerProAsnValValLeuLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu 369
Db AACTCCGAGAATTTGGTCAATTTTGGACGGCCCGGAATCGAGAACTTATTTAAATATTAGAG 228

```



```
Qy 250 TrpAsnProValAlaAhenValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
Db 584 TGAATCCAATGGCAATGTAGGGGCAAGATGTATGGAGTTCCTTCGACCATGGAT 525
Qy 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
Db 524 GTGCTGTGAACATTCATTCATCAAGGGATACATCTCTATTGGTTGGAGCCTTGCCACA 465
Qy 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
Db 464 AGGCCAGTTCCTCGGCCCAACATGAGAGAGAAAGGAGTGGTGGTGGAGGATGCCACA 405
Qy 310 AlalysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsn 329
Db 404 GCTAAGGAATGTGCTCTTCAAGGAATCTGAACAGGGTGATGACGCCCACTTAAT 345
Qy 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
Db 344 GGCAATGMAACGGGGCTGCCCATGCNAACGGG--GCTGCCACTGTTGCTGATATTTTC 288
Qy 350 AsnSerProAenValValAsnLeuSerArgThrGlyIleGluAenLeuAlaLysLeuGlu 369
Db 287 AACTCCGAGAATTTGGTCAATTTGAGCAGCCCGGAATCGAGAACTTATTAAAAATTAGAG 228
Qy 370 AspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAla 389
Db 227 AACCGAAGGAACCATGGCTTGTGTGCTCTACGCCCCCATGGTGCCTCAATTTTGCACAGGT 168
Qy 390 MetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysVal 409
Db 167 ATGGAAGCATCCTATGTGTAATTTGGCCGATAGTTACAGGGAGT---GGAAATAAGGTG 111
Qy 410 GlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGly 429
Db 110 GGAATAATTCAGGCGAGATGGCATGAGAGAAATCTCAAAACAAAGAAATTACAGCTCGGA 51
Qy 430 SerPheProThrIleLeuPhePheProLysHisSerSerArgPro 444
Db 50 AGTTTCCAAACATCTTTCTTCCCTAAACACTCATCTCGCCCA 6

RESULT 13
CK263241 968 bp mRNA linear EST 03-AUG-2004
LOCUS EST709319 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION clone POABL30 5' end, mRNA sequence.
ACCESSION CK263241
VERSION CK263241.1 GI:39820219
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST709320
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..968
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POABL30"
/tissue_type="abiotic stress treated leaf and root tissue"
```

```
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2e-111 Length: 968
Score: 1166.50 Matches: 230
Percent Similarity: 84.7% Conservative: 25
Best Local Similarity: 76.4% Mismatches: 35
Query Match: 47.7% Indels: 11
DB: 7 Gaps: 4

US-10-731-525-8 (1-466) x CK263241 (1-968)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 80 ATGGCTTTGGCTTTTCATCTTTCACTGCAATTCATGGCTCTCTCTCTCTCTCTCTCAT 139

Qy 20 ---GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGly 38
Db 140 CAACAACCCATAGTATGCCAATTTGGGTAAACATTTCAGCTATTGGATCAGCCA----- 190

Qy 39 AlaValAsn-----PheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAla 56
Db 191 -----AACAAATTTCTTGAATTTCTCTCAGAGGGCTTGTCTGTGAAGCCATTATATGCT 244

Qy 57 GluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAla 76
Db 245 GAACCTTAAAGGAATGATGTCCTTCAGCAGCAACCTTTGTTGCTCTCTGAGGTA 304

Qy 77 SerGlu-----ThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAla 94
Db 305 ACCGAGAAAGTAATAGACCGCAGAGGATTTTGAGAACTGGCTAAGGATCTTTGAAAATGCT 364

Qy 95 SerProLeuGluLeuMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114
Db 365 TCACCTCTTGAGATTATGGATAATGCCCTTTGAGAAATTTGGAGATGATATTGGCATTGCT 424

Qy 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 425 TTCACTGGTGGAGAAAGATGTGCTTTGATAGATATGCACATTTACTGGTCGACCATTT 484

Qy 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAla 154
Db 485 AGAGTGTTCAGCCTTGATACCGGGAGGTGAATCCAGAGACCTACCAACTTTTGTATGCT 544

Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174
Db 545 GTCGAGAAGCACTATGCAATTCACATCGAATACATGTTTCTCTGATTTCTGTGAAGTTCCAG 604
```


175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194
 Db GCATTAGTAAGGACCAAGGCTCTTTCTCATTTTACGAGGATGGCCACCAAGATGCTGC 664

195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
 Db CGTGTATTAGGAAGATTAGGCCCTTGGAGAGACCTCAAGAGTTGGGTGCTGGATCAG 724

215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
 Db GGGCAAGAAAGATCATGCTCCCTGGAACTCGATCTGAAGTTCAGTTGTCAGTAGAC 784

235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254
 Db CCCTCTTCGAGGATTTGGATGGTGGCTCTGGAAGCTTGGTGAAGTGGAAACCCGGTGGCT 844

255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
 Db AATGTAGAGGCAAGGACATATGAACCTTCTAGTGCATGAATGTTCCCGTGAACCTCT 904

275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
 Db TTGCATTTGAAGGATATGCTCTATTGGATGTGAACCTTGCACCCAGGCGGTCTACCT 964

295 Gly 295
 Db GGG 967

RESULT 14
 CB340603/c 704 bp mRNA linear EST 14-MAR-2003
 LOCUS CA23BI011VR All Cabernet Sauvignon Leaf - CA23BI Vitis vinifera
 DEFINITION cDNA clone CA23BI011VR All 3', mRNA sequence.

ACCESSION CB340603
 VERSION CB340603.1 GI:28961185
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 704)
 Jones da Silva, F., Lim, H., Iandolino, A., Baek, J., Leslie, A., Xu, J.,
 Gomes, K., Walker, M.A. and Cook, D.R.
 Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 Unpublished (2003)
 Contact: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: GCCAACGCAATGGCTCTAG.
 Location/Qualifiers
 1..704
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /culturvar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CA23BI011VR All"
 /sex="Hermaprodite"
 /dev stage="Mid-season leaf material, collected July 25,
 2001."
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Leaf - CA23BI"
 /note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
 SfiI; CA23BI is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on July 25, 2001, in
 Napa Valley, California, and represent leaves in
 mid-season development. These leaves were verified to be
 infected with the bacterial pathogen, Xylella fastidiosa,

based on a diagnostic assay using PCR and Xylella-specific
 primer pairs. The plants were asymptomatic at the time of
 collection, but later developed symptoms. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AACAGTGGTATCAGCAGAGTGGCCATTACGCCGGG-3' and
 5'-ATTCTAGAGCGGAGCGGCCGACATG-dt(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
 Alignment Scores: 2,91e-111 Length: 704
 Pred. No.: 1163.00 Matches: 209
 Score: 97.4% Conservations: 14
 Best Local Similarity: 91.3% Mismatches: 6
 Query Match: 47.6% Indels: 0
 DB: 6 Gaps: 0

US-10-731-525-8 (1-466) x CB340603 (1-704)

Qy 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGly 117
 Db GAAATTTATGATGAAGCCCTTGCAGAAATTTGGCAATGACATGCCATTGCTTCAGTGGT 644

Qy 118 AlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPhe 137
 Db GCTGAAGATATTGCTTTGATTGAATATGCACGCTTAACCTGGTCTGCTCGTTAGGGTATTC 584

Qy 138 SerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLys 157
 Db AGCCTGGACACTGGAAGGTTGAAACCCGGAAACATATCAATTTCTTGACACGGTTGAGAAA 524

Qy 158 HisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuVal 177
 Db CACTATGGCATCCGCATTGAATACATGTTTCCAGATGCTGTTGAAGTTTCAAGGCTTGCTG 464

Qy 178 ArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgValArg 197
 Db AGGAGCAAGGAGCTGTTCTCTTTTACGAGGATGGCCACCCAGGAGTGTCTCCGCTGTGAGG 404

Qy 198 LysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArg 217
 Db AAGGTGAGACCTTGGAGAGGCACTGAAAGGTTTGGTCCCTGGATCACAGGCGCAAGA 344

Qy 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237
 Db AAAGATCAGTCTCTGCTGACGGCTGAAGTTTCTGTTGTCAGGTGGATCCAGCCTTC 284

Qy 238 GluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLys 257
 Db GAGGGGTGGATGTTGGGTGGTGGCAGCTGGTGAATGGAACCCAGTGGCAATGTGCAG 224

Qy 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAla 277
 Db GGCATGGACATCTGGAATTTCTCTCCGTGCTATGAATGTGCTGTGAATTCATTGCACTCG 164

Qy 278 LysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHis 297
 Db AAAGGGTATATTTCAATTTGGGTGGTGGAGCCATGCACTAGGCTGTATTACTTGCCACAT 104

Qy 298 GluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLys 317
 Db GAGAGAGAGGAGATGGTGGTGGAGAGATGCCAAGGCCAAGGAGTGTGGGCTTCATAAA 44

Qy 318 GlyAsnValLysGlnGlnLysGluGlu 326
 Db GGAACCTCAAAACAAAAA 17

RESULT 15
 CK287268 856 bp mRNA linear EST 02-AUG-2004
 LOCUS EST749990 Nicotiana benthamiana mixed tissue cDNA library,
 DEFINITION

normalized, full-length Nicotiana benthamiana cDNA clone NMB793 5' end, mRNA sequence.

ACCESSION

CK287268

VERSION

CK287268.1 GI:39863647

KEYWORDS

EST.

SOURCE

Nicotiana benthamiana

ORGANISM

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 856)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Skaskawicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr. Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..856
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NMB793"
 /tissue types="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Alignment Scores:

Pred No.:	2,11e-110	Length:	856
Score:	1156.00	Matches:	226
Percent Similarity:	88.0%	Conservative:	31
Best Local Similarity:	77.4%	Mismatches:	25
Query Match:	47.3%	Indels:	10
DB:	7	Gaps:	5

US-10-731-525-8 (1-466) x CK287268 (1-856)

Qy	4	AlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGluProLysLeu	23
Db	1	GCATTCATAGCTCTTTGTCC---CCTTCTCTCTTAT-----GAACATCAAGTG	48
Qy	24	ProGlnIleGlySerIleArgIleSerGluArgPro-----IleGlyGlyAlaValAsn	41
Db	49	TCCCAATTGGGACCTTTTCAGCCATTGGATAGGCTCAAAATGCTGTCGAAGGGCTTTTGAAC	108
Qy	42	PheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArgLys	61
Db	109	-----ACTCGAGGCGTTCGGCAGTGAACCATTTGAACGCTGAGCCCTAAGAGGAAT	159
Qy	62	AspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGlu---ThrLys	80
Db	160	GATTCATAGTTGCTCTTCGACGAACCTCGTGGCTCCTGAAGTAGCAGAGAAACTGAG	219
Qy	81	GluGluAspPheGluGlnIleAlaSerAspLeuAsnAlaSerProLeuGluIleMet	100

Db	220	GAAGAGGACTTCAGAAATTTGGCTAAGGAGCTTCAACATGCTTCCCTCTTGAGATTATG	279
Qy	101	AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp	120
Db	280	GACAAAGCCCTTGAGAAATTTGGGATACGTTGCTATTGCTTTTCAGTGTGCTGAAGAT	339
Qy	121	ValAlaIleGluIleGluValAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp	140
Db	340	GTAGCTTTGATTGAGTATGCACATTTAACTGGTCGACCATTCAGAGTTTTCAGCCITGAT	399
Qy	141	ThrGlyArgLeuAsnProGluThrTyrrGlnLeuPheAspAlaValGluLysHisTyrrGly	160
Db	400	ACTGGGAGGTTGAACCGGAGACCTACCAATATTATTGATAGTGTGGAGAGCATTATGGC	459
Qy	161	IleArgIleGluTyrrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys	180
Db	460	ATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAAGTTTCAGGCTTAGTAGCCAAA	519
Qy	181	GlyLeuPheSerPheTyrrGluAspGlyHisGlnGluCysCysArgValArgLysValArg	200
Db	520	GGGCTTTTCTTCTTCTATGAAGACGGCCACCAAGAGTGTGCCGTGTAAGAGGTTAGG	579
Qy	201	ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln	220
Db	580	CCCTTAAGAGAGCCCTCAAGGCTTACGTGCTGGATCACAGGCAACGTAAGATCAA	639
Qy	221	SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet	240
Db	640	TCCCTTGGAACTCGATCAGAAATTCGCTGTTTTCAGGTGAGCCCTTCTTTGAAGGTTG	699
Qy	241	AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaValValLysGlyHisAsp	260
Db	700	GATGGCGGTGCTGGTAGCTTGGTGAAGTGAACCTGTGGCTAACGTAGACGGAAGAT	759
Qy	261	IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr	280
Db	760	ATTGGAACTTCTTCGTGCAATGAATGTTCCGGTTAACTCATTCATTACATTACAAAGTTAT	819
Qy	281	ValSerIleGlyCysGluProCysThrArgProVal	292
Db	820	GTCTCCATTGGATGCGAAACCTTGCACAAGGCCGGTC	855

Search completed: February 21, 2006, 01:51:34

Job time : 5585 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 20, 2006, 22:26:17 ; Search time 322 Seconds
(without alignments)
2572.496 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPSTPSSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/abssWEB.spool/US10731525/runat.17022006.164901.185/app_query.fasta_1
-DB=Issued PatentsEPOOL -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs06p -USER=US10731525 @CGN 1.1.290 @runat.17022006.164901.185 -NCPU=6
-ICPU=3 -NO KMAP -NEG SCORES=0 -WAIT -DEBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2445	100.0	1629	3	US-09-720-318A-7
2	1923	78.7	1795	3	US-09-720-318A-5
3	1645.5	67.3	1827	3	US-09-720-318A-9
4	1295.5	53.0	1210	3	US-09-720-318A-3
5	1262.5	51.6	1215	3	US-09-720-318A-1
6	811.5	33.2	783	3	US-09-328-352-1063
7	745.5	30.5	810	3	US-09-252-991A-13944
8	546.5	22.4	6063	3	US-09-902-540-807
9	540	22.1	723	3	US-09-902-540-7978

10	407.5	16.7	756	3	US-09-540-236-1037	Sequence 1037, Ap
11	405.5	16.6	269223	3	US-09-596-002-41	Sequence 41, Appl
12	307.5	12.6	747	3	US-09-134-001C-1462	Sequence 1462, Ap
13	307.5	12.6	4055	3	US-09-710-279-3357	Sequence 3357, Ap
14	307.5	12.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
15	307.5	12.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
16	286	11.7	660	3	US-09-252-991A-13837	Sequence 13837, A
17	266.5	10.9	1240	3	US-09-962-357-7	Sequence 7, Appli
18	250.5	10.2	468	3	US-09-710-279-307	Sequence 307, App
19	209.5	8.6	747	3	US-09-543-681A-2919	Sequence 2919, App
20	202	8.3	783	3	US-09-489-039A-2784	Sequence 2784, Ap
21	168.5	6.9	640681	3	US-09-790-988-1	Sequence 1, Appli
22	157.5	6.4	1664976	3	US-08-916-421B-1	Sequence 1, Appli
23	157.5	6.4	1664976	3	US-09-692-570-1	Sequence 10, Appl
24	156	6.4	2413	3	US-09-613-182-10	Sequence 1, Appli
25	147	6.0	2640	3	US-09-962-357-1	Sequence 2, Appli
26	145	5.9	1547	2	US-08-557-122A-2	Sequence 2, Appli
27	145	5.9	1547	3	US-09-262-666-2	Sequence 2, Appli
28	140	5.7	2170	3	US-09-807-258-9	Sequence 9, Appli
29	139.5	5.7	1101	3	US-08-984-919A-13	Sequence 13, Appl
30	139.5	5.7	1101	3	US-08-984-919A-14	Sequence 14, Appl
31	139.5	5.7	1401	3	US-08-984-919A-32	Sequence 32, Appl
32	139.5	5.7	1401	3	US-08-984-919A-34	Sequence 34, Appl
33	139.5	5.7	1410	3	US-08-984-919A-54	Sequence 54, Appl
34	139.5	5.7	1410	3	US-08-984-919A-56	Sequence 56, Appl
35	139.5	5.7	1466	3	US-08-984-919A-10	Sequence 10, Appl
36	139.5	5.7	1488	3	US-08-984-919A-49	Sequence 49, Appl
37	139.5	5.7	1488	3	US-08-984-919A-50	Sequence 50, Appl
38	139.5	5.7	1518	3	US-10-396-869A-7	Sequence 7, Appli
39	139.5	5.7	1569	3	US-09-368-588-1	Sequence 1, Appli
40	139.5	5.7	1875	3	US-08-984-919A-46	Sequence 46, Appl
41	139.5	5.7	1875	3	US-08-984-919A-48	Sequence 48, Appl
42	139.5	5.7	4040	3	US-09-999-833A-458	Sequence 458, App
43	139.5	5.7	4040	3	US-10-020-445A-458	Sequence 458, App
44	138.5	5.7	1107	3	US-08-781-420-13	Sequence 13, Appl
45	138.5	5.7	1107	3	US-08-781-420-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-09-720-318A-7
; Sequence 7, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-318A-7

Alignment Scores:
Pred. No.: 3 86e-270 Length: 1629
Score: 2445.00 Matches: 466
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-731-525-8 (1-466) x US-09-720-318A-7 (1-1629)

QY 1 MetAlaLeuAlaPheThrSerSerFileSerAlaProThrSerThrPheProSerGlu 20
|||||

```
Db      32  ATGCCCTCGCTTTCACTCTTCAATTTCCGACCAACTTCCACCTTCCCATCATCGAA  91
Qy      21  ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal  40
Db      92  CCCAAATTCGCAAAATGGGTCAATTAGGATTTGGAGAGGCCCATTTGGAGGCGCGTT  151
Qy      41  AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg  60
Db      152  AATTTCAATTAATCTCAAGACGGAGCTTGTAAAGCCCTTAACGCCGCACTCCACGC  211
Qy      61  LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys  80
Db      212  AAGGATTCATGTTCTCTCGCAGCAACAACCATGCTTGTCTTCTGAGACGAA  271
Qy      81  GluGluAspPheGluGlnIleAlaSerAspLeuAsnAlaSerProLeuGluIleMet  100
Db      272  GAGGAAGATTTGAACAGATAGCCAGTGATCTGCACAAATGCTTCACTCTTGAATCATG  331
Qy      101  AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp  120
Db      332  GATAGACCCCTCGACAAATTCGGACACGACATGCTATTGCTTCAGTGTGCTGAAGAT  391
Qy      121  ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp  140
Db      392  GTTGCTTTGATTGATGATGGAATTCACGGTTCGACCCCTTAGGGTTTTTCAGTTTGGAC  451
Qy      141  ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly  160
Db      452  ACTGGAGACTGAAACCCAGAACTTATCACTTTTGTATCGGTTGAGAAAGCATATTGGA  511
Qy      161  IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys  180
Db      512  ATTCGCATTAGTACATGTTCCCTCGATGCTGTTGAGGTTTACGCAATGTTGGTGAAGTAAG  571
Qy      181  GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg  200
Db      572  GGGTTATCTTCTTACGAGGATGGCCACCAAGAGTGTTCAGGGTGAGAAAGTTCAGG  631
Qy      201  ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln  220
Db      632  CCTTTAAGGAGGCGCCTTAAGGCTCTCAGACATGGATACTTGTTCAGAGGAAAGACAG  691
Qy      221  SerProGlyThrArgSerGluLeuProValValGlnValAspProAlaPheGluGlyMet  240
Db      692  TCACCTGTGATAGTCTGAAATACCGGTTGTTTCAAGTTGATCCGGCTTTTGGAGGAATG  751
Qy      241  AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp  260
Db      752  GATGTTGAATTTGGAAGCTTGTGAGTGAAGTGAACCTGTGTGCAATGTGAAGGGCCATGAC  811
Qy      261  IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr  280
Db      812  ATATGGAACCTTCCTTAGCACCATGAATGTGCTGTGAATTCCTTTCATGCAAAAGGATAT  871
Qy      281  ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu  300
Db      872  GTTTCCATTTGGTGTGAGCCCTGCACTAGGCGCTGTTTTACCTGGGCAACATGAAGGGAA  931
Qy      301  GlyArgTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal  320
Db      932  GGGAGTGTGTGTGGGAGGATGCCAAGCTAAGGNAATGTGTCTTCACAAAGGAATGTA  991
Qy      321  LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly  340
Db      992  AAGCAGCAGAAAGAGGAGCATGTTAATGCAAAATGGGCTATCCCAATCCCATGCAAAATGGT  1051
Qy      341  AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr  360
Db      1052  GATGCTACCATGTCCTGCAATTTTCAACAGCCCGCAATAGTAGTTAACTTGAGCAGGACT  1111
Qy      361  GlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyr  380
Db      1112  GGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGGAAGGAACCATGGCTTGTGTGCTTAT  1171
```

```
Qy      381  AlaProTyrCysProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLys  400
Db      1172  GCACCAATGGTGCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGACAAG  1231
Qy      401  LeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGlu  420
Db      1232  TTAGCAGGGTCAACAGCGATGAAGGTTGAAAATTTAGAGCAGATGGAGAACAGAAAGAA  1291
Qy      421  PheAlaLysSerGlnLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHis  440
Db      1292  TTTGCAAAAGAGTGAATGCAATTTGGGAAGCTTCCCTACGATATTATTTTCCCAAGCAT  1351
Qy      441  SerSerArgProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAla  460
Db      1352  TCCTCTCGGCCACCAATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCTTTGATGGCA  1411
Qy      461  PheValAsnAlaLeuArg  466
Db      1412  TTTGTAATGCGCTTAAGA  1429

RESULT 2
US-09-720-318A-5
; Sequence 5, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-318A-5

Alignment Scores:
Pred. No.:      3,64e-210      Length:      1795
Score:          1923.00        Matches:      378
Percent Similarity: 87.2%      Conservatives: 37
Best Local Similarity: 79.4%   Mismatches:   45
Query Match:     78.7%        Indels:       16
DB:              3           Gaps:          7

US-10-731-525-8 (1-466) x US-09-720-318A-5 (1-1795)
Qy      1  MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr  15
Db      56  CTTGCCGCTTCCACTACTTCTCTCTCAGTCAGCAGCAGCAGCGCTCGAGTCT  115
Qy      16  Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg  30
Db      116  TTCTTCTCGCGCCTTGGATCTTCATCGGAGCGCTAAAGCTCGCAAAATTTGGTTCCTTCGG  175
Qy      31  IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu  50
Db      176  TTTCCGAGAGGCCCTCAAGTTTCGTCTGCTGTTTAAATTTAACTCAAGACGCTCCTCG  235
Qy      51  ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr  70
Db      236  GTGAGGCCACTCAATCGCAACCGCAACCGGAATGATTCTGTGTTCTCTTTCGACGA---  292
Qy      71  ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp  90
Db      293  ACTATCGTTGCTCTCGAGTTGAGAGGAGGAAGAAATTTTGAGCAATTTAGCCAAAGAC  352
Qy      91  LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp  110
```

```
Db 353 CTTGAAATTCATCTCTCTGAGATTATGATGAAGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleaIaIleAlaPheSerGlyAlaGluaspValaLeuileGluTyrAlaLysLeuThr 130
Db 413 ATGGTATTCCTTAGTGTGCTGAAGATGTTGCTTTGATTGATGATGCACATTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuaspThrGlyArgLeuaspProGluThrTyrGln 150
Db 473 GGTGCGACCTACAGAGTGTAGTCTTGACACTGGGAGACTGAACCCAGAAACCTACAAA 532
Qy 151 LeuPheaspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProaspAla 170
Db 533 TTTTTCACCGCTGTGAGAGACATTATGSAATTTCATATTGAGTACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluaspGlyHis 190
Db 593 GTTGAGGTTCAGGATAGTAAGACTAAGGGCTCTTCTCATTTTACGAGGATGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTCTGTAGTAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTyrPheThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGATTACTGGACAGAGAAAGACCACTCTCTGGTACTAGTCTGAATCCCTATT 772
Qy 231 ValGlnValaspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyr 250
Db 773 GTCAGGTGATCTCTGTTTGTAGGAGCTGATGGTGGAAATTCGACGCTGTGGAATGG 832
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTGCAATGTTAATGTTCTAGACATATGGAATTCCTTAGGACCATGAATGT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CCTGTAATTCATTGTCATTCCTCAAGGATATGTTTCGATTGGCTGTGAGCCATGCACAAG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTyrPheGluaspAlaLysAla 310
Db 953 CCGTTTTTACCCGACACATGAAAGAGAGAGGTTGGTGGGAGGATGCAAGGCC 1012
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluaspValAsnGly 330
Db 1013 AAGGAGTGTGCTTTCACAAAGTAAATTTGAACAGAGAGATGCTGCCCATTAATGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProaspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAAATGGCTCTGCC---ACTGTTGTGCATTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluasp 370
Db 1124 TCCAGATGTGTGCTAGCTGAGCAGTCCGGGATGAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTyrLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAACACTGGCTGTGTGCTCTATGCACCATGGTCCGCTTCGTGAGGCTATG 1243
Qy 391 GluGluSerTyrValaspLeuAlaaspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1244 GAGGAGTCGTATGTTGATCTGCGCAGAGAGTTCAGCAAGGTCA---GGAGTGAAGGTTGCA 1300
Qy 411 LysPheArgAlaaspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1301 AAATTTCAGACCCATGAGAGCAGAGAGGAAATGCAAGAGTGAATCTGCAAGTTGGGAAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1361 TTCCCAACATATCTTCTTCCCAAGCACTCTTCTCAACCA---ATTAAATACCTTCA 1417
Qy 451 GluLysArgaspValaspSerLeuMetAlaPheValAsnAlaLeuArg 466
```

```
Db 1418 GAAAAGAGAGATGTTGATTCAATGACGGCATTCGTGAATGCCTTACGG 1465
RESULT 3
US-09-720-318A-9
; Sequence 9, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-318A-9
Alignment Scores:
Pred. No.: 2,62e-178 Length: 1827
Score: 1645.50 Matches: 327
Percent Similarity: 79.2% Conservative: 49
Best Local Similarity: 68.8% Mismatches: 78
Query Match: 67.3% Indels: 21
DB: Gaps: 8
```

US-10-731-525-8 (1-466) x US-09-720-318A-9 (1-1827)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 89 ATGGCTTCGCT---ACTGCTTCCATCTCG-----TCGCACCTCCATCGCCTGCGCGAT 139
Qy 21 ProLysLeuProGlnIleGlySerIleArg-----IleSerGluArgProIleGlyGly 38
Db 140 CTCAAAGCCCGAGGATTGGAGCCGTGGAGCAGCAGGTGGCCGTGTTCTTCGCGGCGCTG 199
Qy 39 AlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluPro 58
Db 200 CCGGCAACGGCGCCCAAGGGCCAGCGCGAGGGCGTGGCGCGCTGTGCGCGCGGAG 259
Qy 59 ProArgLysAspSerIleValProLeuAlaAlaThrIleValAla----- 74
Db 260 CCAGCGAGGAAGCAGTGTGCGGCTCCGCGGCTCGTCCGCGGTGGCGCGGAGGAG 319
Qy 75 SerAlaSerGluThrLysGluGluaspPheGluGlnIleAlaSerAspLeuaspAsnAla 94
Db 320 GAGGCATCTGCGTGGCGCGCGTGGACTACGAGGCCCTGGCGAGGAGCTGTGGGCGCG 379
Qy 95 SerProLeuGluIleMetaspArgAlaLeuaspLysPheGlyAsnaspIleAlaIleAla 114
Db 380 TCGCCCTCGAGATCATGATCGTGGCTCGACATGTCGGCTCCGAAATCGCATCGCC 439
Qy 115 PheSerGlyAlaGluaspValaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 440 TTCAGTGGTCCGAGAGACGTGGCCCTCATCGAATACGCGAAACTGATGACGCCCTTC 499
Qy 135 ArgValPheSerLeuaspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheaspAla 154
Db 500 AGGCTGTTTCAGCTTCACACTGGCGACCTGAACCCAGAGACATACGACTTTCGACAG 559
Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProaspAlaValGluValGln 174
Db 560 GTGGAGAGCACTATGTTTCCACATCGAGTACATGTTCCCTGAGGCCAGGAGTGCAA 619
Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluaspGlyHisGlnGluCys 194
Db 620 GACCTTGTGAGGAGCAAGGGCTCTTCTTTCTACGAGAGCGGACCGACCGAGGAGTCTGC 679
```

```
QY 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
Db 680 AGGTGAGGAAGTTTCGGCCCTTTCAGAGAGGCCCTTCAGGGCCTCAAGGCTCGATCACC 739
QY 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
Db 740 GGCAGAGGAAGGATCAGTCCCTGGCCACAGAGCCAGCATCCCTGTGTTCAGTTGAT 799
QY 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254
Db 800 CCGTCTTTGAAGGGCTGGATGGTGGAGCCGAGTTCATCAAGTGAACCCCTGTGGCT 859
QY 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db 860 AATGTGATGCAAGGATATCTGACCTTCTCAGHACCATGATGTCCTCTGTGACACC 919
QY 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 920 CTGATGCTCAAGGCTAGCTCCTCATTGGGTGGAGCCGTGCACAGGCCCGTGTGGCG 979
QY 295 GlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGluCysGly 314
Db 980 GGCAGACACGAGAGGAGGAGTGTGTGGAGGACGCCACGGCCAAAGGAGTGGCGC 1039
QY 315 LeuHisLysGlyAsnValLys-----GlnGlnLysGluGluAspValAsnGlyAsn 331
Db 1040 CTGCACAAAGGTACATCGACAGGAAGGTTCAGGCCACCAAGTCGCGCTCAACGGCAC 1099
QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db 1100 GGCTCGCCGAGCCAGCGCC-----CCAGACATCTTCCAGAGC 1138
QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1139 CAGGCCCATCGTCACCTCACCCTCCCGGGATCGAGAACCCTCTCGCGCTCGAAGACCG 1198
QY 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db 1199 GCCAGCCGCGTGCCTACCGCTCTACGCTCCCTGGTGGCCATCTCCAGGCCATGGAG 1258
QY 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1259 CGGTCTCTACGTTCAGCTGCGCAGAGCTCAGCGGCTCA--GGCATCAAGGTGGCCAG 1315
QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1316 TTCCGCGCGGACGCGAGCAGAGCCATTTCGCGCAGCGGAGCTGCAACTACAGAGCTTC 1375
QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1376 CCGACGATCTCTCTGTTCCCGCGCGCCACCGTGAGGCC--ATCAAGTACCCCTCCGAG 1432
QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1433 AAGAGGACGCTCCAGTCCCTCTCGCTTCCTGTAACAGCCCTCAGA 1477
```

RESULT 4

```
US-09-720-318A-3
; Sequence 3, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1210
```

```
; TYPE: DNA
; ORGANISM: Impatiens balsamia
US-09-720-318A-3

Alignment Scores:
Pred. No.: 1,95e-138 Length: 1210
Score: 1295.50 Matches: 239
Percent Similarity: 88.1% Conservative: 28
Best Local Similarity: 78.9% Mismatches: 29
Query Match: 53.0% Indels: 7
DB: Gaps: 3
```

US-10-731-525-8 (1-466) x US-09-720-318A-3 (1-1210)

```
QY 165 TyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184
Db 9 TACATGTTCCCTGATGCAATTGTAGTACAGGATTAGTAAACCAAGACACTGTTCTCT 68
QY 185 PheTyrGluAspGlyHisGlnGluCysArgValArgLysValArgProLeuArgArg 204
Db 69 TTCACGAGACGACATCAAGAGTGTCTCCGCGTCAGAAAAGTGAGGCCACTGAGCGT 128
QY 205 AlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThr 224
Db 129 GCTCTCAAGGGTCTCCGCGCTTCGATCACCGGGCAAGAAAAGACCAAGTCGCCGGGAACG 188
QY 225 ArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyLe 244
Db 189 AGATCGAGAGATCCAGTCGTCGTCAGGTGATGCCCTCTTTTGGAGGATGTTGTTGGAGAG 248
QY 245 GlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPhe 264
Db 249 GGTAGCTGTGTGAAGTGGATCCGCTGGCTAATGTAGATGCTGCTGATGTATGAAATTC 308
QY 265 LeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGly 284
Db 309 CTCGAGCTATGAATGCTGCTTAATGCACTTCATAGCCAGGGTTATGTCTCGAATGGG 368
QY 285 CysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTip 304
Db 369 TGCGAACCGTGCACCGCGGTGTACCTGGGCAACATCAGAGAGAAAGCAGGTGTGG 428
QY 305 TrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLys 324
Db 429 TGGAGAGATCTCGCGCTAAGAGTGTGGCTACATAAAGAAATATAAAG----- 479
QY 325 GluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThr 344
Db 480 -----GATCCCAATGGGAATGGGTTGCTCAAGCTGAGGGAGGAGGAAGAACTGTTACG 533
QY 345 ValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsn 364
Db 534 GATGCTGATATTTTGAATCCAGAATGTGGTGACACTGAGTAGAAGCGGATTTGAGAAT 593
QY 365 LeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCys 384
Db 594 CTGTGGAACCTTCAGGAGAGGAAGACCATGATCGTGTGCTGTATGCACCTTGTGTC 653
QY 385 ProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySer 404
Db 654 CAGTTCTGCCAGGATATGGAAAAATCATACTTGGAAATGGCTGAAAAAGCTGGCGGTGAGC 713
QY 405 ---ThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLys 423
Db 714 GTGCGTGTGTGAAGGTAGGAAATCCGGGCAGATGTCAGAAAAGAGATTGTCTCAC 773
QY 424 SerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHisSerSerArg 443
Db 774 CAAGAATTGCAGCTCGCGGAGCTTTCCAAACAATACTCTCTTCCCAAAACACTCATCTAAA 833
QY 444 ProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsn 463
Db 834 ---GCCATCAAGTACCCATCTGAGAAAAGGACGCTGAGTCAATTGTTGGCTTTTGTGAAC 890
```


Score: 546.50 Matches: 118
 Percent Similarity: 59.3% Conservative: 45
 Best Local Similarity: 42.9% Mismatches: 81
 Query Match: 22.4% Indels: 31
 DB: 3 Gaps: 9

US-10-731-525-8 (1-466) x US-09-902-540-807 (1-6063)

Qy 46 GlnArgSerLeuValProValAsnAlaGluProArgLysAspSerIle 64
 Db AAGAGCGGGCTTTCTTCTCGAGTACCATGCTCGCCCGCAGACCCAG----- 3824

Qy 65 ValProLeuAlaThrThrIleValAlaSerAlaSerGluThrLysGluGluAspPhe 84
 Db -----GCTCGGCGCTCTCGCGGAGCGCTGAAGTGAAGGACGCGCC----- 3779

Qy 85 GluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeu 104
 Db -----GCGGAGACCTC-----CTCGCTGGACCGGCGC----- 3749

Qy 105 AsplysPheGlyAsnAspIleAlaIleAlaPheSer---GlyAlaGluAspValAlaLeu 123
 Db -----CGCTTCGTCGCGCGCGCCATCGCTCCAGCTTCGCGTGGAGGACATGCTCTC 3692

Qy 124 IleGluThrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArg 143
 Db ATCGACCTGGCGGCGGACGATGCGCCAGCGCTGCGCTCTTCACGCTCGACACCGGACGC 3632

Qy 144 LeuAsnProGluThrThrGlnLeuPheAspAlaValGluLysHisThrGlyIleArgIle 163
 Db CTGCCCCGAGAGCGTACGAACTCATGGAGTGTGCTAAGCGCTACGCGCTACCGGTG 3572

Qy 164 GluThrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPhe 183
 Db GAGAGCTACTTCCCGAGCGCGCGCTGAGGCGTGGAGTCCAGACGCGTACTTTC 3512

Qy 184 SerPheThrGlu-----AspGlyHisGlnGluCysCysArgValArgLysValArgPro 201
 Db TCCTTCGCGCAGAGCGCTAGAGGACGCAAGCGCTGCTGCGCCATCCGCAAGTGGAGCCC 3452

Qy 202 LeuArgAlaLeuLysGlyLeuArgAlaThrIleThrGlnArgLysAspGlnSer 221
 Db CTGTGCGCGCGCTGCGGCGTCAACAGCGTGTGAGCGGATTCGCGCGTGGAGCGTCC 3392

Qy 222 ProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAsp 241
 Db -----GGGCTGTCTCAAGCTCAACCGCTGGCCACGTTAGAGTGGACAGCGGCATC 3347

Qy 242 GlyGlyIleGlySerLeuValLysThrAsnProValAlaAsnValLysGlyHisAspIle 261
 Db -----GGGCTGTCTCAAGCTCAACCGCTGGCCACGTTAGAGTGGACAGCGGCATC 3299

Qy 262 TrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyThrVal 281
 Db TGGGCTTACCTCGCGGCGAAGCGTGGCTTCAACCGCTTCCATGACCGTGGCTACCCG 3239

Qy 282 SerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGly 301
 Db TCATCGGCTGTGGCGCTGCACGCGCGCGTGAACCCCTACGAGGACGCGCGCGGCG 3179

Qy 302 ArgThrTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis 316
 Db CGCTGTGTGGAGTCCGCGGAGAACCGCGAGTGGCGGCTCCAC 3134

RESULT 9

US-09-902-540-7978
 ; Sequence 7978, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 7978
 ; LENGTH: 723
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 ; US-09-902-540-7978

Alignment Scores:
 Pred. No.: 4,2e-52 Length: 723
 Score: 540.00 Matches: 114
 Percent Similarity: 59.9% Conservative: 43
 Best Local Similarity: 43.5% Mismatches: 75
 Query Match: 22.1% Indels: 30
 DB: 3 Gaps: 8

US-10-731-525-8 (1-466) x US-09-902-540-7978 (1-723)

Qy 58 ProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSer 77
 Db 7 CCCCCGAGCCCGCAG-----GCCTCGCGCTCTCGCGCGAAGCGCT 48

Qy 78 GluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeu 97
 Db 49 GAAGTGAAGACGCCCC-----GCCAGGACCTC-----CTC 81

Qy 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116
 Db 82 GCCTGGACCGAGCGC-----CGCTTCGTCGCGCGCGCGCCATCGCTCCAGCTTC 132

Qy 117 GlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGlyArgProPheArgVal 136
 Db 133 GCGCTGGAGGACATGCTCTCATCGACCTGGCGCGCCAGCATCGCGCCAGCTCGCGCTC 192

Qy 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu 156
 Db 193 TTCACCTGACACCGAGCGCTGCCCCGAGACGTACGACTCATGAGGTGGTGGCT 252

Qy 157 LysHisThrGlyIleArgIleGluThrMetPheProAspAlaValGluValGlnAlaLeu 176
 Db 253 AAGCGCTACGGCTCACCGTGGAGACGTACTTCCCGAGCGCGCGCGCGGTGG 312

Qy 177 ValArgSerLysGlyLeuPheSerPheThrGlu-----AspGlyHisGlnGluCysCys 194
 Db 313 GAGTCCCAACGCGCTACTTCTCTCCCGCAGAGCTAGAGCGCAGCGGTGTGTC 372

Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
 Db 373 GCCATCCGAAGGTGAGCGCTGTCGCGCGCTCGCGGTCAACAGGCGTGGTGGACG 432

Qy 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
 Db 433 GGATTGGCGCGTGGAGCAGTCCGTC---ACCGCACCGACGTGGCGCGCTTAGAGGTGAC 489

Qy 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysThrAsnProValAla 254
 Db 490 AGCGCGCAT-----GGGCTGTCTCAAGCTCAACCGCTGCGC 525

Qy 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
 Db 526 ACCTGAGGAGCGCGGAGCATCTGGCGCTAGCTCCGGCGGAGAGCGTGCCTCAACGCG 585

Qy 275 LeuHisAlaLysGlyThrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
 Db 586 CTCCATGACCGTGTACCGTCCATCGCTGTCGCGCTGCGCGCGCGCGGTGAAACCC 645

Qy 295 GlyGlnHisGluArgGluGlyArgThrTrpTrpGluAspAlaLysGluCysGly 314
 Db 314

Db 646 TAGGAGACGAGCGCGCGCGCTGGTGGTGGAGTCCGCGGAGAACCCGAGTGGCG 705

Qy 315 LeuHis 316
|||||

Db 706 CTCAC 711

RESULT 10

US-09-540-236-1037

; Sequence 1037, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS

; FILE REFERENCE: 2709-2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 1037

; LENGTH: 756

; TYPE: DNA

; ORGANISM: M.catarrhalis

US-09-540-236-1037

Alignment Scores:

Pred. No.: 7.26e-37 Length: 756

Score: 407.50 Matches: 88

Percent Similarity: 56.3% Conservative: 33

Best Local Similarity: 40.9% Mismatches: 77

Query Match: 16.7% Indels: 17

DB: 3 Gaps: 5

US-10-731-525-8 (1-466) x US-09-540-236-1037 (1-756)

Qy 106 LysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGlu 125

Db 151 AAGTTGCCAGTAGTTTGCAG-----GCCGAGGATATGGTCATCCACGAT 195

Qy 126 TyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsn 145

Db 196 GCACTGGCAATAGCCATTTAAGTACTGAAATATTTATTTTACAGACAGGTAGCTTAAC 255

Qy 146 ProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyr---GlyIleArgIleGlu 164

Db 256 GCCGAAACACTTAAGCTCATCGCATCGTCAGGCAATGCTTATCTTACAAATAATTTTAA 315

Qy 165 TyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184

Db 316 ACTTATGAGCGCACCCGCGCTGGCAGACTATGTGGCAATCATGGCTTAAACGCT 375

Qy 185 PheTyrGluAspGly-----HisGlnGluCysCysArgValArgLysValArgProLeu 202

Db 376 TTTTACGAAAGTGGTGATCTACGCAATTTGCTGCTTTATTCGTAAGTTGAGCCTTTA 435

Qy 203 ArgGlnAlaLeuLysGlyLeuArgAlaTrrPileThrGlyGlnArgLysAspGlnSerPro 222

Db 436 AATCGTGTCTTGGTCGATGCTGGTGAGCGGCAACCCGCGTGAAGTCAAGTCAAGTC 495

Qy 223 GlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAspGly 242

Db 496 ---ACTGCGACGAGCTGAATTTGGCAGAACACAGATTGGCG----- 534

Qy 243 GlyIleGlySerLeuValLysTrrPAsnProValAlaAsnValLysGlyHisAspIleTrrp 262

Db 535 -----CGTGATTTGCCAAATACAAATTTTGTATTTGGCAAGACGAGCGATGTGG 588

Qy 263 AsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer 282

Db 589 GCATATTTTGGCAAAATAATACCGTTTATGAGCTGTATACACCAAGGCTATCTTCC 648

Qy 283 IleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlyArg 302

Db 649 ATTGGCTGTAGCCTTGTACCATGCTGTCAAGCAAGCGAGGATATCCGTCGAGGACGC 708

Qy 303 TrpTrrPTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLys 317
|||||

Db 709 TGGTGGTGGAGCATAAAGATAATAAGAGTGGCGTTTACACAA 753

RESULT 11

US-09-596-002-41

; Sequence 41, Application US/09596002

; Patent No. 6632636

; GENERAL INFORMATION:

; APPLICANT: Lagace, Robert, E.

; APPLICANT: Berg, Kim, L.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME

; FILE REFERENCE: PM-0008-4 US

; CURRENT APPLICATION NUMBER: US/09/596,002

; CURRENT FILING DATE: 2000-06-16

; PRIOR APPLICATION NUMBER: 60/140,121

; PRIOR FILING DATE: 1999-06-18

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PERL Program

; SEQ ID NO 41

; LENGTH: 269223

; TYPE: DNA

; ORGANISM: Moraxella catarrhalis

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte template ID No. 6632636 41

; PUBLICATION INFORMATION:

US-09-596-002-41

Alignment Scores:

Pred. No.: 1.93e-32 Length: 269223

Score: 405.50 Matches: 89

Percent Similarity: 56.7% Conservative: 33

Best Local Similarity: 41.4% Mismatches: 76

Query Match: 16.6% Indels: 17

DB: 3 Gaps: 6

US-10-731-525-8 (1-466) x US-09-596-002-41 (1-269223)

Qy 106 LysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGlu 125

Db 232582 AAGTTGCCAGTAGTTTGCAG-----GCCGAGGATATGGTCATCCACGAT 233026

Qy 126 TyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsn 145

Db 233027 GCATTGGCAATAGCCAGTTAAGTACTGAAATATTTATTTTACAGACAGGTAGCTTAAC 233086

Qy 146 ProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyr---GlyIleArgIleGlu 164

Db 233087 GCCGAAACACTTAAGCTCATCGCATCGTCAGGCAATCATGGCTTAAACGCT 233146

Qy 165 TyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184

Db 233147 ACTTATGAGCGCACCCGCGCTGGCAGACTATGTGGCAATCATGGCTTAAACGCT 233206

Qy 185 PheTyrGluAspGly-----HisGlnGluCysCysArgValArgLysValArgProLeu 202

Db 233207 TTTTACGAAAGTGGTGATCTACGCAATTTGCTGCTTTATTCGTAAGTTGAGCCTTTA 233266

Qy 203 ArgGlnAlaLeuLysGlyLeuArgAlaTrrPileThrGlyGlnArgLysAspGlnSerPro 222

Db 233267 AATCGTGTCTTGGTCGATGCTGGTGAGCGGCAACCCGCGTGAAGTCAAGTCAGTC 233326

Qy 223 GlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAspGly 242

Db 233327 ---ACTGCGACGAGCTGAATTTGGCAGAACACAGAT-----TTGACGCGT 233368

Qy 243 GlyIleGlySerLeuValLysTrrPAsnProValAlaAsnValLysGlyHisAspIleTrrp 262

Db 233369 GGTATT-----GCCAAATACAAATTTTGTATTTGGCAAGACGAGTGTATGG 233419

QY 263 AsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer 282
Db 233420 GCATATATTTGACCAAAATATACCGTTTAAATGAGCTGATCACCAAGGCTATCCTTCC 233479
QY 283 IleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArg 302
Db 233480 ATTGGCTGTGAGCCTTGTACCATGCTGTCAAGCAAGCGGAGGATATCCGTGAGGAGCG 233539
QY 303 TrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLys 317
Db 233540 TGGTGGTGGGAGGATTAAGATTAAGAGTGGGTTTACACAAA 233584
RESULT 12
US-09-134-001C-1462
; Sequence 1462, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1462
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1462
Alignment Scores:
Pred. No.: 2,13e-25 Length: 747
Score: 307.50 Matches: 75
Percent Similarity: 51.4% Conservative: 39
Best Local Similarity: 33.8% Mismatches: 91
Query Match: 12.6% Indels: 17
DB: 3 Gaps: 8
US-10-731-525-8 (1-466) x US-09-134-001C-1462 (1-747)
QY 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116
Db 106 GAGATTTTAAATGGGCATATCAAACTATGAAATGATATGTTTATCTTCAGATTTT 165
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
Db 166 GGTGCTGAGAGTATGTTTGTGATTGACTTCTCAAAATTAACCCGACGACAAAT 225
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaVal--- 155
Db 226 GTATTTTATAGACACTGATTACATTTTCAAGAAACATATGACTTAATAGATAGGGTAA 285
QY 156 GluLysHisTyrGlyIleArgIleGluTyrMetPheProasp---AlaValGluValGln 174
Db 286 GATAAATATCCGAATACGATTAATAATGAAAAACCCGAACTTACGTAGAGGAA 345
QY 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194
Db 346 GGTGAGAAATATAATCCTGCTTTA-----TGAAGAATGATCCTAACCAATGTTGC 396
QY 195 ArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
Db 397 TACATACGCAAGATTAACACCACTAGAGACGTATATCTGGTCTGTAGCTGTGATATCA 456
QY 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValasp 234
Db 457 GGTCTTAGAGCAGACATACCA---ACACAGACATACAAATTTCAATTAACAAGAT 513
QY 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254

Db 514 GAAAGATTTAAGTCAATTAATA-----GTGTGTCCTTAATCTATTGGACAGAA----- 561
QY 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db 562 -----GAAGAATATGCTCTTATATACGTGATAAGGACTTACCTTTAATGAA 609
QY 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 610 TTACATGATCAAAATTTATCAAGTATTTGGTTCATTCCATGTATACATCACCCTGATTGAT 669
QY 295 GlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGlyCysGly 314
Db 670 TCTAATGATTCAGTCTGCTGCTGCT-----TGTCCAATTTCTAGTAAGACTGAATGCGGA 723
QY 315 LeuHis 316
Db 724 TTACAT 729
RESULT 13
US-09-710-279-3357/c
; Sequence 3357, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3357
; LENGTH: 4055
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3357
Alignment Scores:
Pred. No.: 3,43e-24 Length: 4055
Score: 307.50 Matches: 75
Percent Similarity: 51.4% Conservative: 39
Best Local Similarity: 33.8% Mismatches: 91
Query Match: 12.6% Indels: 17
DB: 3 Gaps: 8
US-10-731-525-8 (1-466) x US-09-710-279-3357 (1-4055)
QY 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116
Db 2981 GAGATTTTAAATGGGCATATCAAACTATGAAATGATATGTTTATCTTCAGATTTT 2922
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
Db 2921 GGTGCTGAGAGTATGTTTGTGATTGACTTCAAAATTAACCCGACGACAAAT 2862
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaVal--- 155
Db 2861 GTATTTTATAGACACTGATTACATTTTCAAGAAACATATGACTTAATAGATAGGGTAA 2802
QY 156 GluLysHisTyrGlyIleArgIleGluTyrMetPheProasp---AlaValGluValGln 174
Db 2801 GATAAATATCCGAATACGATTAATAATGAAAAACCCGAACTTACGTAGAGGAA 2742
QY 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194
Db 2741 GGTGAGAAATATAATCCTGCTTTA-----TGAAGAATGATCCTAACCAATGTTGC 2691
QY 195 ArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214

; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:

Pred. No.:	3,37e-19	Length:	4411529
Score:	307.50	Matches:	98
Percent Similarity:	44.6%	Conservative:	38
Best Local Similarity:	32.1%	Mismatches:	124
Query Match:	12.6%	Indels:	45
DB:	3	Gaps:	12

US-10-731-525-8 (1-466) x US-09-103-840A-1 (1-4411529)

Qy	34	ArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysPro	53
Db	2686275	CGACCG--GGTGGTGGCAACTTCGTCAAACACCGCAGCGAGGTGAACGCTTCGGCGCA	2686331
Qy	54	ValAsnAlaGluProArgLysAspSerIleValProLeuAlaAlaThrIleVal	73
Db	2686332	GTG---GGTCATCCGGCGCGAGGAGGACGACCTCGCATGAGCGCGGAGACCAACC	2686382
Qy	74	AlaSerAlaSerGluThrLysGluGluAspPheGluGlnIle	89
Db	2686383	-----AGGTGACCGAACCCGCAACTAGTGAGTGGCGCGCGGAGCTGCC	2686430
Qy	90	AspLeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsn	109
Db	2686431	GAATCGACGGCGCCACCGCCACCGACATGTTGGCTGGACCGACGAACCTTCGGC	2686487
Qy	110	AspIleAla	112
Db	2686488	GACATCGCGCGCGCGCGCGCGGTGAGCGGACATCGCGGTGACACCGTGCACACTAC	2686547
Qy	113	IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGlyArg	132
Db	2686548	GTAGTCTCTTCAACATGGCTGATGCGGTGCTGGTGGATGCGCGCCCAAGGTGCGACCG	2686607
Qy	133	ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe	152
Db	2686608	GGCGTACCGGTCTCTTCTTGATACCGGTACCACTTCGTCGAAACATCGGCACACAGA	2686667
Qy	153	AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAsp	171
Db	2686668	GATCGCATCGAGTCCGCTCTATGACGTCGGGTGCTCAATGTCACCTCGGAGCACACAGTG	2686727
Qy	172	GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln	191
Db	2686728	GCCGAGCAGGACGCACTGGGCAAGGACTTGTGCGC-----CGCAACCCCAT---	2686778
Qy	192	GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla	211
Db	2686779	GAATGTCGCGGTGGCAAGGTGCTTCCCTGGGCAAGACGCTGCTGGTGTACTCGCG	2686838
Qy	212	TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal	231
Db	2686839	TGGGTGACCGGGTACGGCGGGTCGATGCACCG---ACCGGGCCAATGCCCGCTGCTC	2686895
Qy	232	GlnValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTrpAsn	251
Db	2686896	AGCTTCGATGAGACGTTCAAA-----CTAGTGAAGGTCAAC	2686931
Qy	252	ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro	271
Db	2686932	CCGCTGGCGCGGTGGACCGCAAGATGTCCAGAAATACATGCCAGAACGACGTGCTG	2686991
Qy	272	ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro	291
Db	2686992	GTAAATCCGCTTGGCGGAAGGCTATCCGTCGATCGTTGCTCGCTCGTGACAGCCAAA	2687051
Qy	292	ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLys	311
Db	2687052	CCCGCCGAAGGCGCGACCCGCGACGCGACG-----TGGCAGGGGCTGGCCAGACC	2687105

Qy	312	GluCysGlyLeuHis	316
Db	2687106	GAATGCGGGTTGCAC	2687120

Search completed: February 20, 2006, 23:57:58
Job time : 3412 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 20, 2006, 22:59:55 ; Search time 1395 Seconds
(without alignments)
2762.387 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPTSTPSSSE.....KYPSEKRDVDSLMFVNALR 466

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10731525/runat.17022006.164909.374/app_query.fasta_1
-DB=Published Applications_NA_Main -QFWT=fascap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -DOCALIGN=200 -THR SCORES=pct
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -ALIGN=15 -MAXLEN=2000000000 -HOST=abs06p
-USER=US10731525 @CGN 1.1 1549 @runat.17022006.164909.374 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -JSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Published Applications NA Main.*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2445	100.0	1629	7	US-10-731-525-7
2	2445	100.0	1629	7	US-10-762-049-7
3	1951	79.8	1730	7	US-10-425-114-7777
4	1951	79.8	2774	7	US-10-424-599-134742
5	1923	78.7	1795	7	US-10-731-525-5
6	1923	78.7	1795	7	US-10-762-049-5
7	1811.5	74.1	1398	3	US-09-938-842A-498

8	1811.5	74.1	1398	3	US-09-938-842A-498
9	1738	71.1	1365	3	US-09-938-842A-2305
10	1738	71.1	1365	3	US-09-938-842A-2305
11	1645.5	67.3	1827	7	US-10-731-525-9
12	1645.5	67.3	1827	7	US-10-762-049-9
13	1645.5	67.3	1840	8	US-10-739-930-2784
14	1639.5	67.1	1880	7	US-10-437-963-2751
15	1597.5	65.3	1380	7	US-10-169-667A-1
16	1509.5	61.7	1553	8	US-10-425-115-83222
17	1444.5	59.1	1850	8	US-10-425-115-83219
18	1422.5	58.2	1927	8	US-10-425-115-83223
19	1295.5	53.0	1210	7	US-10-731-525-3
20	1295.5	53.0	1210	7	US-10-762-049-3
21	1262.5	51.6	1215	7	US-10-731-525-1
22	1262.5	51.6	1215	7	US-10-762-049-1
23	1241.5	50.8	1183	7	US-10-425-114-15300
24	1203.5	49.2	1156	7	US-10-425-114-12185
25	1105	45.2	711	7	US-10-424-599-135400
26	1039	42.5	1194	8	US-10-425-115-83221
27	999.5	40.9	1294	8	US-10-425-115-83224
28	914	37.4	549	8	US-10-425-115-147208
29	871.5	35.6	985	7	US-10-767-701-10835
30	754	30.8	796	7	US-10-424-599-82919
31	731.5	29.3	705	6	US-10-369-493-37482
32	716.5	29.3	471	3	US-09-732-627A-2057
33	697.5	28.5	585	7	US-10-021-323-3831
34	679	27.8	623	7	US-10-021-323-11704
35	571	23.4	742	8	US-10-425-115-12975
36	539	22.0	690	6	US-10-369-493-42894
37	537	22.0	330	7	US-10-767-701-17535
38	520	21.3	588	7	US-10-021-323-3903
39	509	20.8	552	7	US-10-021-323-11747
40	472.5	19.3	56609	9	US-10-915-740A-7
41	472.5	19.3	2242716	9	US-10-915-740A-1068
42	454.5	18.6	599	7	US-10-424-599-69970
43	449.5	18.4	675	6	US-10-369-493-43366
44	439	18.0	514	3	US-09-770-961-142
45	405.5	16.6	269223	7	US-10-672-787-41

ALIGNMENTS

RESULT 1
US-10-731-525-7
; Sequence 7, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; PRIOR FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-10-731-525-7

Alignment Scores:			
Pred. No.:	3.84e-271	Length:	1629
Score:	2445.00	Matches:	466
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

```
US-10-731-525-8 (1-466) x US-10-731-525-7 (1-1629)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 32 ATGGCCCTCGCTTTCACTTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGAA 91
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal 40
Db 92 CCCAAACTTCGCGAATTTGGGTCAATTAGGATTTCCGAGAGGCCCATTTGAGGGCCGCTT 151
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
Db 152 AATTTCAATTTATCTCAAGACCGGCTTGGTAAGCCCGTTAAACGCCGACCTCCACGC 211
Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerGluThrLys 80
Db 212 AAGGATTCATGTTCTCTCGACCAACCACTCGTTGCTTCTGCTTCTGAGACGAAA 271
Qy 81 GluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100
Db 272 GAGGAAGATTTTGAACAGATAGACGAGTATCTGACCAATGCTTCCACCTCTTGAATCATG 331
Qy 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120
Db 332 GATAGAGCCCTCGACAAATTCGCGCAACGACATACCTATTGCTTTCAGTGGTGTGAAGAT 391
Qy 121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140
Db 392 GTTGTCTTTGATTGATGATCGGAAATTCAGCGGTGCGACCTTTTAGGGTTTTCAGTTTGGAC 451
Qy 141 ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly 160
Db 452 ACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTGCGGTGAGAACATATTGGA 511
Qy 161 IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys 180
Db 512 ATTCGCATTGATGATACATGTTCCCTGATGCTGCTGAGGTTCAGGCTTGGTGAGGAGTAAG 571
Qy 181 GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg 200
Db 572 GGGTATTCTCTTACGAGGATGGGCACCAAGAGTGTTCAGGGGTGAGAAAGGTGAGG 631
Qy 201 ProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln 220
Db 632 CCTTTAAGAGGGGCCCTTAAGGGTCTCAGACATGGATACTGTGTGAGAGAAAGACGAG 691
Qy 221 SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet 240
Db 692 TCACCTGGTACTAGGTCTGAAATACCGGTTGTTTCAGGTTGATCCGGCTTTTGAGGGAATG 751
Qy 241 AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp 260
Db 752 GATGGTGAATTTGGAAGCTTGGTGAAGTGAACCTGTTGCAAAATGTGAAGGGCCATGAC 811
Qy 261 IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr 280
Db 812 ATATGGAACTTCTCTAGGACCATGAATGTGCTGTGAATTCCTTGATGCAAAAGGATAT 871
Qy 281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu 300
Db 872 GTTTCCATTGGGTGTGAGCCCTGACATAGGCTGTGTTTACCTGGGCAACATGAAGGGAA 931
Qy 301 GlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320
Db 932 GGGAGGTGGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGTTCTTCACAAAAGGAAATGTA 991
Qy 321 LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly 340
Db 992 AAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCATGCAAAATGGT 1051
Qy 341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr 360
Db 1052 GATGCTACCACTGTGCTGACATTTTCAACAGCCGGAATGTAGTTAACTTTCAGGAGGACT 1111

RESULT 2
US-10-762-049-7
; Sequence 7, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-7

Alignment Scores:
Pred. No.: 3 84e-271 Length: 1629
Score: 2445.00 Matches: 466
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-731-525-8 (1-466) x US-10-762-049-7 (1-1629)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 32 ATGGCCCTCGCTTTCACTTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGAA 91
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal 40
Db 92 CCCAAACTTCGCGAATTTGGGTCAATTAGGATTTCCGAGAGGCCCATTTGAGGGCCGCTT 151
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
Db 152 AATTTCAATTTATCTCAAGACCGGCTTGGTAAGCCCGTTAAACGCCGACCTCCACGC 211
Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerGluThrLys 80
Db 212 AAGGATTCATGTTCTCTCGACCAACCACTCGTTGCTTCTGCTTCTGAGACGAAA 271
```



```
Db 457 GGTGCGACCTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAAA 516
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 517 TTTTGTGATCGGTGAGAGCATATATGGAATTCACATTGAGTACATGTTCCCTGATGCG 576
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 577 GTTGAGGTTTCAGGCATTAGTAAGAACTAAAGGGCTCTTCATTTTACGAGGATGGGCAT 636
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 637 CAAGAGTGCTGCCGAGTAAGAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCATA 696
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 697 GCATGGATCATCTGCACAGAGAAAGATCAGTCTCTCGTACTAGGTCAGAAATCCCTGTT 756
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 757 GTCCAGGTTGATCTGTTTGTAGGGHACTGATGGTGAATTTGGCAGCCCTGGTGAAGTGG 816
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgTrpMetAsnVal 270
Db 817 AACCCGGTTCGAATGTTAATGGTCTAGACATATGGAGCTTCCTTAGGACCATGGATGTT 876
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 877 CCTGTAAATTCATTGTCATCTCCAGGATATGTTTCGATTTGGCTGTGAGCCGTGCACAAG 936
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 937 CCAGTTTTACCGGTCAACATGAAGAAGAGAAAGGTGGTGGTGGAGAGATGCCAAAGCC 996
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db 997 AAGGAGTGTGCTTTCACAAAGGTAATAATAAACAATCAAGATGTGCGCCAGCTTAATGGA 1056
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsn 350
Db 1057 AATGGGCGCTCCCAA-----GCAAAATGGATCTGCC---ACTGTGCTGACATTTTCAAC 1107
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1108 TCCAGGATGGTGGCTTGGACAGGTCTGGGATTCAGAAATTTGGCAAAATTTGGAGAAC 1167
Qy 371 ArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1168 CGAAAGAACCCCTGGCTTGTGTTGCTCTATGCAACCATGGTGTGCGCTTCTGTGAGGCTATG 1227
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1228 GAGGAATCATATGTTGATTTGGCAAAAGATAGCTGGGTCA---GGAGTGAAGGTGTGCA 1284
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1285 AAATTCAGCGGATGAGACCAAGGAATATCAAGACTGAACCTGCACTGCAGTGGGAGC 1344
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1345 TTCCCCACAATACTACTCTTCCCCCAACACACTCTTCTCAACCA---ATTAAGTACCCTTGG 1401
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1402 GAAAGAGAGATGTTGATTCATTGACGCGCATTCGTGAATGCTTTACGG 1449
```

RESULT 4

```
US-10-424-599-134742
; Sequence 134742, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
```

```
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21 (53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 134742
; LENGTH: 2774
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92682C.1
US-10-424-599-134742
```

```
Alignment Scores:      8.53e-214      Length:      2774
Pred. No.:             1951.00        Matches:     382
Score:                 88.2%          Conservative: 38
Best Local Similarity: 80.3%          Mismatches:  40
Query Match:           79.8%          Indels:      16
DB:                    7              Gaps:        7
```

```
US-10-731-525-8 (1-466) x US-10-424-599-134742 (1-2774)

Qy 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThr 15
Db 106 ATGGCTCTCGCCGTTTCCACTACTTCTCTTCAGACGCTGCTGCGCTTCAGTCC 165
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 166 TTCTTCTCTCGTCTCGGATCTTCTCCGACGCCAAAGCTCTCAAAATCGGTGCTTTCGG 225
Qy 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgSerIleu 50
Db 226 TTTCCGAGAGGTGCGTCTGTTCTGTTGTGTTCAATGTAACCAACGACGCTCCTTG 285
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 286 GTGAGGCCACTCAACGCCGAAACCGCAACGGAAGATTTCTATTGTTCTCTTTCAGCA--- 342
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 343 ACTATCGTGTCTCTGAGGTTGAGAGGAAGAGGAGATTTTGAGCAAAATAGCGAAAGAC 402
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 403 CTTGAAAATGCATCTCCTCTTTGAAATTTATGATAGGGCCCTCGAGAAAATTTGGGAACGAC 462
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 463 ATCGCTATTGCATTTAGTGTGCTGGAAGATGTTGCTTTGATTGAGTATGCACATTTGACG 522
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 523 GGTGACCCCTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAAA 582
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 583 TTTTTCGATCGGTTGAGAAGCATTTATGGAATTCACATTTGAGTACATGTTCCCTGATGCG 642
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 643 GTTGAGGTTGAGGCATTAGTAAGAACTAAAGGGCTCTTCTCATTTTACGAGGATGGGCAT 702
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 703 CAAGAGTGCTGCCGAGTAAGAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCATA 762
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 763 GCATGGATCACTGGACAGAGAAAAGATCAGTCTCCTGCTACTAGGTCTGAAAATCCCTGTT 822
```

```
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 823 GTCCAGGTGATCTGTTTGGAGGACTGGATGGTGAATGGCAGCCTGGTGAAGTGG 882
Qy 251 AsnProValAlaAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgTrpMetAsnVal 270
Db 883 AACCCGGTTGCAATGTTAATGGTCTAGACATATGGAGCTTCCTTAGGACCATGGATGTT 942
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 943 CCTGTAATTCATTTGCAATCCCAAGGATATGTTTCGATTTGGCTGTGAGCCGTGCACAAAG 1002
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 1003 CCAGTTTTACCGGTCAACATGAAGAGAGGAAGGTGGTGGTGGAGGATGCCAAAGCC 1062
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1063 AAGGAGTGTGGTCTTCAAAAGGTAATATAAAACATGAAGATGCTGCCAGCTTAATGGA 1122
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1123 AATGGGGCTCCCA-----GCMAATGGATCTGCC---ACTGTGCTGACATTTTCAAC 1173
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1174 TCCAGGATGTGGTCAGCTTGGAGCAGGTCTGGGATTGAGAATTTGGCAAAATTTGGAGAAC 1233
Qy 371 ArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1234 CGAAAGAACCTGGCTTGTGTGCTATGCAACATGCTGGTGGCCGCTTCTGTCAAGCTATG 1293
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1294 GAGGAATCATATGTTGATTTGGCAAAAGATTAGCTGGTCA---GGAGTGAAGTTGCA 1350
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1351 AAATTCAGAGCGATGAGACCAAGAGGAATATGCAAAAGACTGAACCTGCAGTTGGGAAGC 1410
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1411 TTCCCCACAACTACTCTTCCCCAAACACTCTTCTCAACCA---ATTAAGTACCCCTTCG 1467
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1468 GAAAGAGAGATGTTGATTCATTGACGGCATTCTGTAATGCTTTACGG 1515
```

RESULT 5

```
US-10-731-525-5
; Sequence 5, Application US/10731525
; Publication No. US2004012140A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; PRIOR FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Glycine max
US-10-731-525-5
Alignment Scores:
```

```
Pred. No.: 7,64e-211 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservative: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: Gaps: 7

US-10-731-525-8 (1-466) x US-10-731-525-5 (1-1795)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTTGCCGTTTCCACTACTTCTTCTTCTAGCTGCAGCAGCAGCAGCGGTCCAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCATCGACGCTAAAGCTCCGCAAAATGGTTCCTTTGG 175
Qy 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeu 50
Db 176 TTTCCGGAGAGGCTCAAGTTTCTGCTGTTGTTAAATTTAACTCAAGACGCTCCTCG 235
Qy 51 ValLysProValAsnAlaGluProAspLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCCACTCAATGCCAACCCGCAACGGAATGTTCTGTGTTCTCTTTCAGCA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCGTTGCTCTCGAGTTGAGAGGAGAAAGAGATTTTTCAGCAATTTAGCGAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAAATTCATCTCTCTTGAGATTATGATGAAGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 413 ATCGCTATTGCTCTTGTAGTGTCTGAAGATGTTGCTTTGATTGAGTATGCACATTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 473 GGTGACGCTTACAGAGTGTGTTAGTCTTGACACCTGGGAGACTGAACCCAGAAACCTACAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 533 TTTTGTGAGCGCTTTCAGAGCAATTTATGGAATTCATATTCAGTACATGTTCCCTGTATCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 593 GTTGAGGTTTCAGGCAATTAGTAAGAACTAAGGGGCTCTTCTCATTTTACGAGGATGGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGTGTAGAGTAAGAAAGGTGAGGCCCTTCGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGATTACTGGACAGAGAAAGACCAGTCTCTGGTACTAGGTCTGAAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 773 GTCCAGGTGATCCTGTTTGTAGGGACTGGATGGTGGAAATTTGGAGCCCTGGGAAGTGG 832
Qy 251 AsnProValAlaAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTGCAATGTTAATGGTCTAGACATATGGNACTTCTTAGGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CCTGTAATTCATTTGCAATCCCAAGGATATGTTTCGATTTGGCTGTGAGCCATGCACAAAG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 953 CCGGTTTATACCCGGCAACATGAAGAGAGGAAGGTGGTGGTGGGAGGATGCGCAAGCC 1012
```

```
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db 1013 AAGAGTGTGGTCTTCAAAAGGTAATTTGAAACAGAGAAATGCTGCCAGCTTAATGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAAATGGCTCTGCC---ACTGTTGCTGCATTTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGGTTCAGCTTTCAGCAGGTCCGGGATTCAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTropLeuValLeuTyrAlaProTropCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAACACTGGCTTGTTGCTCTATGCACCATGGTCCGCTTCTGTGAGCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1244 GAGGAGTCGTATGTTGATCTGGCAGAGAAAGTTAGCAAGGTCA---GGAGTGAAGGTTGCA 1300
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLysGlySer 430
Db 1301 AAATTCAGAGCCGATGAGAGCAGAGAGAAATATGCAAAAGAGTGAACCTGCAGTTGGGAAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1361 TTCCCCAACATCTCTCTTCCCCAGACACTCTTTCTCAACCA---ATTAGTACCCCTCA 1417
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1418 GAAAGAGAGATGTTGATTCATTGACGCGCATTCGTGAATGCCTTACGG 1465

RESULT 6
US-10-762-049-5
; Sequence 5, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-5

Alignment Scores:
Pred. No.: 7,64e-211 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservative: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: 7 Gaps: 7

US-10-731-525-8 (1-466) x US-10-762-049-5 (1-1795)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTTCGGGTTTCCACTACTTCTTCTCTCTCAGCTGCAGCAGCAGCAGCGCTCGAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCTATCGACGCTAAAGCTCGCAAAATGGTTCCTTCGG 175
```

```
Qy 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
Db 176 TTTCGGAGAGGGCTCAAGTTTCGTCTGGTGTGTAAATTTAACTCAAGACGCTCCTCG 235
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCCACTCAATCCGAACCGCAACGGAATGATTTCTGTGTCTCTTGCAGCA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCGTTGCTCCTCGAGGTGAGAGGAGGAGAGAGATTTTGGACAAATTTAGCAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAAATTCATCTCTCTTGAGATATGATAAGGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 413 ATCGCTATTGCTTTAGTGTGCTGAAGATGTTGCTTGTGATTTAGTATGCACATTTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 473 GGTGACCCCTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAACCCAGAAACCTACAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 533 TTTTGTGACGCTGTTGAGAGCATTTATGGAATTCATATTGAGTACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 593 GTTGAGGTTTCAGGCATTTAGTAAGAACTAAGAGGGCTCTCTCTCATTTTACGAGGATGGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGTGTAGAGTAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTropIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGNTTACTGGACAGAGAAAGACCACTCTCTCTGTTACTAGGTCTGAAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyr 250
Db 773 GTCCAGGTTGATCCTGTTTGTGAGGACTGGATGGTGGAAATTTGGCAGCCTGGTGAAGTGG 832
Qy 251 AsnProValAlaLeuValLysGlyHisAspIleTropAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTGCAATGTTAATGGTCTAGACATATGGAACTTCCTTTAGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CTTGTAATTCATTGCAATTCGAGGATATGTTTCGATTTGGCTGTGAGCCATGCACAAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTropTropGluAspAlaLysAla 310
Db 953 CCGGTTTTACCGGACAACATGAAAGAGAGGAGTGGTGGTGGGAGGATGCCAAAGCC 1012
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db 1013 AAGGAGTGTGGTCTTCAAAAGGTAATTTGAAACAGAGAAATGCTGCCAGCTTAATGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAAATGGCTCTGCC---ACTGTTGCTGCATTTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGGTTCAGCTTTCAGCAGGTCCGGGATTCAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTropLeuValLeuTyrAlaProTropCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAACACTGGCTTGTTGCTCTATGCACCATGGTCCGCTTCTGTGAGCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
```


Db 1351 AAGAGAGATGTTGAGTCTTGGACTTCGTTCTTGAATCTTGTCCGA 1395

RESULT 8

US-09-938-842A-498

; Sequence 498, Application US/09938842A

; Publication No. US20040009476A9

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 498

; LENGTH: 1398

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-498

Alignment Scores:

Pred. No.: 3.89e-198 Length: 1398

Score: 1811.50 Matches: 351

Percent Similarity: 84.0% Conservative: 48

Best Local Similarity: 73.9% Mismatches: 57

Query Match: 74.1% Indels: 19

DB: 3 Gaps: 7

US-10-731-525-8 (1-466) x US-09-938-842A-498 (1-1398)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19

Db 1 ATGGCAATGCTGTAATATGTTCTCTTCGTCCTTCGGGATCATATAAATCTCGTTTC 60

Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35

Db 61 GGTGTTTCATTTGGAGCCAAAGATTTCGCAATTTGGTTGAGGTTATTGGATCGTGT 120

Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54

Db 121 CATGTTGCTCTGTCCTCTGTAATCTATCTGGGAGCGATCATCATCTGTTAAACCTTTA 180

Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74

Db 181 AACGCTGAACCAAGACAAAGGATTCATGATTTCTCTTCGCGCAACAATGGTAGAGAA 240

Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

Db 241 ATTGCAGAGGAAGTTGAAGTGGTTCGAGATTGAGGATTTTGAAGAGCTTTGCTAAAGATTA 300

Qy 92 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111

Db 301 GAGAAATGCTTACCTCTTGAGATTATGACAAAGCTCTTGAGAAATACGGGACGATATC 360

Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131

Db 361 GCCATTGCATTTAGTGGTGCAGAGATGTTCTCTTATTGAGTAGCGCTCATTTGACTGGG 420

Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151

Db 421 AGGCCATTTAGATATTATTTAGTTGGATACAGGAGGTTGAATCTCTGAGACGATATCGGTTT 480

Qy 152 PheAspAlaValGluLysHieTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171

Db 481 TTCGATGCGGTGGAGAAGCACATATGGATTAGGATTAGTATATGTTTCTCCTGATTCGTT 540

Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191

Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATTTCTCTTTTATGAGGATGGTCAATCAG 600

Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLysValAla 211

Db 601 GAGTGTTCGCGTGTTCGAAGGTGAGACCTTTGAGCGTGCTCTCAAGGTTTAAAGGCT 660

Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231

Db 661 TGGATTACTGTCAGAGGAAGATCAATCTCCGGGACAAGTCTGAGATTCCGGTTGTT 720

Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251

Db 721 CAGGTTGATCCGGTGTTCGAAGGTTTGGATGGTGGAGTTTGTAGTTTGGTGAAGTGAAT 780

Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271

Db 781 CCGGTTGCGAATGTTGAAGGAATGATGTTTGAACCTCTTGAGGACTATGGATGTTCCG 840

Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291

Db 841 GTTAACACATTTGCATCGCGCAGGATATATATCGATTGATGTGAGCCTTGACGAAAGCG 900

Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLys 311

Db 901 GTTTTACCGGTCAGCACGAGAGAGGAGATGGTGGTGGGAAGATGCTTAAAGCCCAAG 960

Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331

Db 961 GAATGTGGATCTCACAAAGGAATGTCTCAA---GAAACCTCCGATGATGTAAAGTGAAC 1017

Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351

Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056

Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGlnAsnLeuAlaLysLeuGluAspArg 371

Db 1057 GAGAAATCTTGATCTTTGAGCAGCGCAGGAGATTGAGAATTTGATGAAGTTGGAGAACCGT 1116

Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391

Db 1117 AAGAGCCTTGATGATGTCGTCGTTTATGCTCCGTCGTCGTCGTCGTCGTCGTCGTCG 1176

Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411

Db 1177 GCATCGTATGATGAATCGCGGATAAATTTGCTGGAAGT---GGGATTAGGTTGCCAAA 1233

Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431

Db 1234 TTCAGACGATGTCGTCAGACGAGAGGATTTGCTTAAGCAGGAATTTGCGAGCTCGTAGCTTC 1293

Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTrpSerSerGlu 451

Db 1294 CCTACCATTCCTGTTTCCCTTAAGACTCATCGAGACG---ATCAAGTATCCGTCGAG 1350

Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466

Db 1351 AAGAGAGATGTTGATCTTTGACTTCGTTCTTGAATCTTGTCCGA 1395

RESULT 9

US-09-938-842A-2305

; Sequence 2305, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

Qy	233	ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnPro	252
Db	700	GTTGATCCAGTGTTCGAAGGGTTAGATGGCGGTGGAAAGTCTTGTTGAAGATCGAATCCT	759
Qy	253	ValAlaAsnValLysGlyHisAspIleTrpAsnPhelEuArgThrMetAsnValProVal	272
Db	760	TTGGCTTAATGTTCAAGGAGCTGATGTGTGAACCTTCTTGAGAATACTGGATGTTCGGGTG	819
Qy	273	AsnSerLeuHisAlaLysGlyTyValSerIleGlyCysGluProCysThrArgProVal	292
Db	820	AATGCATTGCACGCACACAGGTAATGTCTCATCGGTGTGAGCGGTACTAGGCCGGTG	879
Qy	293	LeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGlu	312
Db	880	CTTCCAGGCCAACATCAGAGAGAAGAGTGGTGGGAAGATGCTAAAGCTAAAGAA	939
Qy	313	CysGlyLeuHisLysGlyAsnValLysGlnLysGluAspValAsnGlyAsnGly	332
Db	940	TGTGGTCTACACAAGGGGAACATC-----AAGGAGGAAGAT-----	975
Qy	333	LeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsnSerPro	352
Db	976	-----GGTGCTGCAGACTCAAAGCCCTGCTGTCMAGAGATATTTGAAGCAAC	1026
Qy	353	AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys	372
Db	1027	AATGTGGTGCATTGAGCAAGGAGGGTGTGAGAATCTTTTGAAGCTAGAGAACCCTAAA	1086
Qy	373	GluProTrpLeuValValLeuTyAlaProTrpCysProTyrcysGlnAlaMetGluGlu	392
Db	1087	GAGCGGTGGTGGTTCGTACTTACGCTCTCTGGTGCCCTTCTGCCAGGCGATGGAAGCA	1146
Qy	393	SerTyValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe	412
Db	1147	TCGTACATCGAATGGCTGAGAAACTTGGCGGAAAA---GGAGTTAAAGTGGCGAAATTC	1203
Qy	413	ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro	432
Db	1204	CGAGCTGACGGTGAGCAGAAGAGATTGTCTAAGCAGAGCTTCAGTTAGGAGCTTCCC	1263
Qy	433	ThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyProSerGluLys	452
Db	1264	ACGATATCTCTCTTCGAAAAAGAGCTCCACGG---GCTATTAAAGTACCTTCAGAGCAT	1320
Qy	453	ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg	466
Db	1321	AGAGATGTTGATTCACCTATGCTGTTGTGAATCTTCTTCGG	1362
 RESULT 10 US-09-938-842A-2305 ; Sequence 2305, Application US/09938842A ; Publication No. US20040009476A9 ; GENERAL INFORMATION: ; APPLICANT: Harper, Jeff ; APPLICANT: Kreps, Joel ; APPLICANT: Wang, Xun ; APPLICANT: Zhu, Tong ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING SUCH GENES ; TITLE OF INVENTION: SAME, AND METHODS OF USE ; FILE REFERENCE: SCRP1300-3 ; CURRENT APPLICATION NUMBER: US/09/938,842A ; PRIORITY FILING DATE: 2001-08-24 ; PRIOR APPLICATION NUMBER: US 60/227,866 ; PRIOR FILING DATE: 2000-08-24 ; PRIOR APPLICATION NUMBER: US 60/264,647 ; PRIOR FILING DATE: 2001-01-16 ; PRIOR APPLICATION NUMBER: US 60/300,111 ; PRIOR FILING DATE: 2001-06-22 ; NUMBER OF SEQ ID NOS: 5379 ; SEQ ID NO 2305 ; LENGTH: 1365 ; TYPE: DNA			

ORGANISM: Arabidopsis thaliana
US-09-938-842A-2305

Alignment Scores:

Pred. No.: 1.14e-189 Length: 1365
Score: 1738.00 Matches: 346
Percent Similarity: 82.5% Conservative: 45
Best Local Similarity: 73.0% Mismatches: 55
Query Match: 71.1% Indels: 28
DB: 3 Gaps: 9

US-10-731-525-8 (1-466) x US-09-938-842A-2305 (1-1365)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
Db 1 ATGGCTTTAGCTGTTACTCTCTTCAACTCGCAATCTCTGGATCGAGTTTTCTCGTTCT 60
Qy 17 --ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 61 GGAGCTTCTTCTGAACTAAAGCTTCAATATGTTGCAATTAGTTATCTGATCGAACC 120
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55
Db 121 -----CATTTGCTCAGAGACGTTACTCTATGAAACCTTTAAAC 159
Qy 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
Db 160 GCTGAGTCACATTCACGAAGCAATCTTGGGTCTACTCGTCTTCAACTCTAATTTGCTCCT 219
Qy 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
Db 220 GAAGTTGAAGAGAAAGAGAGAGAGTTGAAGACTTTTGACCACTTGTCTAAAGACTTGAA 279
Qy 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
Db 280 GATGCTTCTCCACTTGAATCATGATAAGCTCTTCAGAGATTTCGAGACCAATCGCA 339
Qy 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
Db 340 ATTGCTTTTATGAGAGTGAAGATGTTGCAATGATGCAATATGCACGTTTTAACTGGAAG 399
Qy 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
Db 400 CCATTAGGGTTTTTATGTTAGATACAGGGAGATTAAACCTGAAAGCTACAGCTCTTT 459
Qy 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
Db 460 GACGAGTCGAGAGACGAGTACGGGATTCGAATTGAGTACATGTTTCTGATGCAAGTTGAG 519
Qy 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
Db 520 GTTCAAGCTTTAGTGAGAAACAAGGGTTGTTCTTCATTTATGAAGATGGTCAACAGAG 579
Qy 193 CysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTyr 212
Db 580 TGTTCCGCTGACAGAAAGTTAGACCTTTGGCTGCTCTTAAGGGTCTTAAAGCTTGG 639
Qy 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
Db 640 ATTACAGACAGAGAAAGACCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCAG 699
Qy 233 ValAspProAlaPheGluGlyMetAspGlyClyIleGlySerLeuValLysTyrAsnPro 252
Db 700 GTTGATCCAGGTGTTGAAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 759
Qy 253 ValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProVal 272
Db 760 TTGGCTAATGTTGAAGAGCTGATGTGTGGAACTTTCTGAGAACTATGATGTTCCGGTG 819
Qy 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
Db 820 AATGCATTCGACGACAAAGGGTATGTGTCAATCGGTGTGAGCCGTGTACTAGGCCGGTG 879
```

```
Qy 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrTriPduAspAlaLysGlu 312
Db 880 CTTCCAGGCCACATGAGAGAGAGAGAGTGTGTGGAGAGATGCTAAAGCTTAAGAA 939
Qy 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332
Db 940 TGTGGTCTTACACAAAGGGAACATC-----AAGGAGGAAGAT----- 975
Qy 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352
Db 976 -----GGTGTCTCAGACTCAAAAGCCTGCTGCTGCAAGAGATATTTGAAAGCAAC 1026
Qy 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
Db 1027 AATGTGGTTCATTTAGCAAGAGAGGGTGTGAGAACTTTTGAAGCTAGAGAACCGTAAA 1086
Qy 373 GluProTyrPheValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGlu 392
Db 1087 GAGCGCTGTTGGTCTGCTACTTTACGCTCTCTTGTGCGCTTCTGCCAGGCGATGGAAGCA 1146
Qy 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412
Db 1147 TCGTACATCGAATTTGGCTGAGAAACTTGGCGGAAAA---GGAGTTAAAGTGGCGAAATTC 1203
Qy 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
Db 1204 CGAGCTGACGGTGACGAGAGAGGTTGCTTAACAGAGAGCTTCAGTTAGGGAGCTTCCCG 1263
Qy 433 ThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLys 452
Db 1264 ACCGATCTTCTCTTCGAAAAAGAGCTCCACGG---GCTATTAAAGTACCCCTTCAGAGCAT 1320
Qy 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1321 AGAGATGTTGATTCACATCATGTGCTGTTGTGAATCTTCTTCGG 1362
```

RESULT 11

US-10-731-525-9
; Sequence 9, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-731-525-9

Alignment Scores:
Pred. No.: 8.25e-179 Length: 1827
Score: 1645.50 Matches: 327
Percent Similarity: 79.2% Conservative: 49
Best Local Similarity: 68.8% Mismatches: 78
Query Match: 67.3% Indels: 21
DB: 7 Gaps: 8

US-10-731-525-8 (1-466) x US-10-731-525-9 (1-1827)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 89 ATGGCTTCCGCT---ACTGCTTCCATCTCG-----TCGCACCTCCATCGCCCTCGCGAT 139
```



```
Db 380 TCAGCGCTGGAGATCATGGATCGTGGCTCGACATGTTGGCTCCGAAATCGCATCGCC 439
Qy 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 440 TTCAGTGGTCCGAGGACGTGGCCCTCATCGAATACGCGAAACTGACTGGACGCCCTTC 499
Qy 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAla 154
Db 500 AGGTGTTCAGCCTTGACCTGGCGGACTGAACCCAGACATACGAACTCTTCGACAG 559
Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174
Db 560 GTGGAGAGACATATGATCCATCCATCGATGATGTTCCCTGAGGCGCAGGAGTCCAA 619
Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194
Db 620 GACCTTGTGAGGACAAAGGGCTCTTCTCTTCTACGAGGACGACACAGGAGTGTGC 679
Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrrIleThr 214
Db 680 AGGTGTAGGAGGTTCGGCCCTTGAGGAGGCCCTCAAGGGCCCTCAAGGCCCTGGATCACC 739
Qy 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
Db 740 GGGCAGAGGAGGATCAGTCCCTGGCACCCAGCCAGCATCCCTGTTGTTCAGTTGAT 799
Qy 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrrAsnProValAla 254
Db 800 CCGTCTTTGAAGGCTGGATGGTGGAGCGCGGTAGCTTGATCAAGTGGAAACCTGTGGCT 859
Qy 255 AsnValLysGlyHisAspIleTrrAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db 860 AATGTGATGTCAGAGGATATCTGACCTTCTCAGGACCATGGATGTCCTGTGAAACCC 919
Qy 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 920 CTGATGCTCAAGCTAGCTCTCATTTGGGTGGAGCGGTGCACAGGCCGTGTTCGGC 979
Qy 295 GlyGlnHisGluArgGluGlyArgTrrPrrGluAspAlaLysAlaLysGluCysGly 314
Db 980 GGGCAGCAGCAGAGGAGGAGGTGGTGGTGGGAGGACGCCACGGCCAAAGGAGTGGCGC 1039
Qy 315 LeuHisLysGlyAsnValLys-----GlnGlnLysGluGluAspValAsnGlyAsn 331
Db 1040 CTGCACAAACGGTAATCATGACAGGAAGGTGAGGACCCCAAGGTTCGGCGTCAACGGCAAC 1099
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsnSer 351
Db 1100 GGCTCGGCGGAGCCAGCGCC-----CCAGACATCTTCCAGAGC 1138
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1139 CAGGCCATCGTCAACCTCACCCGTCGGGATCGAGAACCTCTCTGGCGCTCGAGAACCGC 1198
Qy 372 LysGluProTrrPrrLeuValLeuTyrAlaProTrrCysProTrrCysGlnAlaMetGlu 391
Db 1199 GCCAGCGCGTGGCTACCGCTCTCTACGCTCCCTGGTGCCCATCTGTCAGGCAATGGAG 1258
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1259 CGGTCTCTAGTGTAGCTGGCCGAGNAGTGGAGCGGTCA--GGCATCAAGTGGCCAG 1315
Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1316 TTCGCGCGGACGGCAGCAGAACCAATTCGCGAGCGGAGCTGCAACTACAGAGCTTC 1375
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTrrProSerGlu 451
Db 1376 CCGACGATCTCTCTGTTCCCGCGCGCACCGTGAAGCCC---ATCAAGTACCCCGTCCGAG 1432
Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
```

```
Db 1433 AAGAGGAGCGTCCAGTCCCTCTCTCGCTTCGTAACAGACCTCAGA 1477
RESULT 13
US-10-739-930-2784
; Sequence 2784, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 2784
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1840)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER4907_2
US-10-739-930-2784
```

```
Alignment Scores:
Pred. No.: 8,34e-179 Length: 1840
Score: 1645.50 Matches: 329
Percent Similarity: 78.2% Conservative: 47
Best Local Similarity: 68.4% Mismatches: 67
Query Match: 67.3% Indels: 39
DB: 8 Gaps: 10
```

US-10-731-525-8 (1-466) x US-10-739-930-2784 (1-1840)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 114 ATGGCATCCGTACCACCTCCATCTCGTCGCACCTCGCCGCTC-----CGCGAC 164
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyAlaVal 40
Db 165 GTCAAAGCTCGAGGATTTGGTGC-----GCGCTGAGGAGGTGGCGCTGGCGCGC 215
Qy 41 AsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsn---AlaGluProPro 59
Db 216 CGT-----GGCCAGCGCGCTCGGGCGGTGCCCCGCTGCGCGCGCGGAGCCTCGC 266
Qy 60 ArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSer-----AlaSer 77
Db 267 AGCGCG-----CCGGTGGCGGCTCCCGCGNCCGCGGTGC-CCNCNCGCGCGC 313
Qy 78 GluThrLysGluGlu-----AspPheGluGlnIleAlaSerAspLeuAspAsnAla 94
Db 314 CCCGTGGCCAGCAGGCGGCGGTGGACTACGAGGCCCTGGCGCGCGAGCTTGAAGGCGCG 373
Qy 95 SerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114
Db 374 TCGCCGTGGAGATCATGGATTCGGCGCGTGGCCATGTTCCGAAATCGCAATCGCATGCC 433
Qy 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 434 TTCAGCGCGCGGAGGACGTGGCGCTGATCGAGTACGCGAAGCTGACGGGCGTCCCTTC 493
Qy 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAla 154
Db 494 CGGCTGTTACGCTGGACACGCGGGCGGCTCAACCGGAGACGTACCACTCTTCGACAGG 553
Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174
Db 554 GTGAGAGACACTACGGCATCCCGCATCGATGATGTTCCCGAGCGCGAGGAGGTGAG 613
Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194
```

```
Db 614 GAGCTGTGGCGACCAAGGCGCTTCTCTCTTACGAGGACGGCCACCAAGGAGTGCTGC 673
Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
Db 674 CGGTGCGCAAGTGCGGCGCTTCCGCGAGCGGCTCAGGGGCTCAGGGGCTGATCACC 733
Qy 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluLeuProValValGlnValAsp 234
Db 734 GGCAGAGGAAGGACCAAGTCCCGCGGACCAAGGCGCAGCATCCCTCTGGTCCAGGTCGAC 793
Qy 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254
Db 794 CCTTCTTCGAAGCGCTGCGAGCGGCGGCTGGTGTGCTGAGTGTGCTCAAGTGGAAACCCCGTGGCC 853
Qy 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db 854 AAGCTGACGCAAGGACATCTGACTTCTCCGACCATGAGCGTCCCTGTCAAGGCC 913
Qy 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 914 CTGCACGCCCGAGGCTGAGTGTCCATCGGGTGGAGCGGTGCACAGGCTGTCTGCGG 973
Qy 295 GlyGlnHisGluArgGlyArgTyrTrpTrpGluAspAlaLysGlyCysGly 314
Db 974 GGGCAGCAGCAGCGGGAAGCGCGTGTGTGGTGGAGGACGCCAAGGCCAAGGAGTGCGGC 1033
Qy 315 LeuHisLysGlyAsnValLysGlnGlnLysGlu----- 325
Db 1034 CTCACAAAGGCAACATCGACAGGACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCC 1093
Qy 326 GluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrVal 345
Db 1094 AGGACTGCCAACGCAAGCGC-----TCGGCGGGCGCC 1126
Qy 346 ProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeu 365
Db 1127 CCGGACATCTTCGAGAGCGCGCGCTGTCTGCTTCAACCGCGCGCGGCTCGAGAACCTG 1186
Qy 366 AlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTrpAlaProTrpCysPro 385
Db 1187 CTGGCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1246
Qy 386 TyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThr 405
Db 1247 TTCTGCCAGGCATGAGGCTCTCTAGTGGAGCTGGCCGAGAGGCTGGCGGGGTCCGGC 1306
Qy 406 GlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGlu 425
Db 1307 GGGCTGAAGGTGGCGCGCTTCCGCGCGAGCGGCGAGCAGAAGCGCTTCCGCGCAGCGCGAG 1366
Qy 426 LeuGlnLeuGlySerPheProThrIleLeuPheProLysHisSerSerArgProThr 445
Db 1367 CTGCAGCTGAGAGCTTCCCGACCGTGTCTGTTCCTCCCGAGCGCGCGCGCGCGCGCGCG 1423
Qy 446 IleLysTyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeu 465
Db 1424 ATCAAGTACCGCTCGGAGAGAGGAGCTGACTCGCTCTCTCGCTTCTGTCACAGGCTC 1483
Qy 466 Arg 466
Db 1484 CGG 1486
```

RESULT 14

```
US-10-437-963-23751
; Sequence 23751, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
```

```
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCES: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 23751
; LENGTH: 1880
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28800C.1
US-10-437-963-23751
```

```
Alignment Scores:
Pred. No.: 4,24e-178 Length: 1880
Score: 1639.50 Matches: 324
Percent Similarity: 78.3% Conservative: 51
Best Local Similarity: 67.6% Mismatches: 71
Query Match: 67.1% Indels: 34
DB: 7 Gaps: 10
```

US-10-731-525-8 (1-466) x US-10-437-963-23751 (1-1880)

```
Qy 10 SerAlaProThrSerThrPheProSerSerGluProLysLeuProGlnIleGlySerIle 29
Db 119 TCTGTCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 174
Qy 30 Arg-----IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArg 47
Db 175 AAGCAGCAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231
Qy 48 ArgSerLeuValLysProValAsnAlaGluProProArgLysAspSerIleValProLeu 67
Db 232 GCGAGGCGGTGGCGCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 282
Qy 68 AlaAlaThrThrIleValAlaSerAlaSerGluThrLysGluGlu----- 82
Db 283 TCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 342
Qy 83 -----AspPheGluGlnIleAlaSerAspLeuAspAsn 93
Db 343 GCGCTCGACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 402
Qy 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113
Db 403 GCGTCTCGCTGTGAGATCATGATCGGCACTCGCCATGTTCGGATCCGACATCGCCATC 462
Qy 114 AlapheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgPro 133
Db 463 GCCTTCAGCGGGCGGAGGACGCTGGCGCTGTATCAGTACGCGAAGCTGACGGGGCGTCCG 522
Qy 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAsp 153
Db 523 TTCGGGTGTTCAGCTTGGACACGGGGCGGCTGAACCCGAGAGACGTACACAGCTGTTCG 582
Qy 154 AlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluVal 173
Db 583 AAGGTGGAGAAGCACTACCGGATCCGATCGATCATGTTCCTCCGACGCGCGCGCGAGGTG 642
Qy 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 193
Db 643 CAGGCGCTGTGTGGCGCGCAAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 702
Qy 194 CysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIle 213
Db 703 TGGCGCGCGCGCAAGGTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 762
Qy 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValGlnVal 233
```



```
Db 943 AAAGCCAAAGGAATCGGGCTCCACAAGGGCAACATCAAGCAA-----GACGAGTCCGTG 996
Qy 329 AsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIle 348
Db 997 -----CCTCCCTCCAGCAACGGCAACGGCACCGCGCGTCCGATCTC 1038
Qy 349 PheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeu 368
Db 1039 TTC-----TCCGTGTTTCTCTGAGCCGGCCAGGCATGGAGAACTTGCTCCGCTG 1089
Qy 369 GluAspArgLysGluProThrLeuValValLeuTyrAlaProThrCysProTyrCysGln 388
Db 1090 GAGAGCCGAGGACTCGTGGCTTGTGCTCTATGCTCCTCGTCCGCTTCTGCCAG 1149
Qy 389 AlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLys 408
Db 1150 GGGATGGAGGAGTCTTCAACGAGTGGCGAGGACTACTCGGCGCGGACTCAGGAGTGAAG 1209
Qy 409 ValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeu 428
Db 1210 GTGGGCAAAATTCAGAGCCGAGGAGACGCGAGAGCTTTCGGGAG---GAGATCGGGCTG 1266
Qy 429 GlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyr 448
Db 1267 GTGAGTTTTCACCATACTCTTCTCCGAAGCACGCGTCAAAACCG---GTCAAGTAC 1323
Qy 449 ProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1324 CCGTCGGAGCAGAGAGAGCTGACTCTCTGTCGCCCTTCGTCAACGCCCTTCGC 1377
```

Search completed: February 20, 2006, 23:30:28

Job time : 1425 secs

THIS PAGE LEFT BLANK